BEST AVAILABLE COPY

ll

(54) Title: PROCESS FOR DESIZING CELLULOSIC FABRIC

(57) Abstract

A process for desizing cellulose-containing fabric comprises treating the fabric with a modified enzyme (enzyme hybrid) which comprises a catalytically active amino acid sequence of an enzyme, particularly a non-cellulolytic enzyme, linked to an amino acid sequence comprising a cellulose-binding domain. A desizing composition suitable for use in the process comprises an enzyme hybrid of the type in question and a wetting agent.

1

PROCESS FOR DESIZING CELLULOSIC FABRIC

5 ## FIELD OF THE INVENTION

The present invention relates to an improved enzymatic process for desizing [i.e. removing "size" (*vide infra*) from] fabric or textile, more particularly cellulose-containing fabric or
10 textile, and to a composition for use in the process.

BACKGROUND OF THE INVENTION

During the weaving of textiles, the threads are exposed to
15 considerable mechanical strain. In order to prevent breaking, they are usually reinforced by coating ("sizing") with a gelatinous substance ("size").

The most common sizing agent is starch in native or modified
20 form. However, other polymeric substances, for example polyvinylalcohol (PVA), polyvinylpyrrolidone (PVP), polyacrylic acid (PAA) or derivatives of cellulose [e.g. carboxymethylcellulose (CMC), hydroxyethylcellulose, hydroxypropylcellulose or methylcellulose] may also be abundant in the size.
25 Small amounts of, e.g., fats or oils may also be added to the size as a lubricant.

As a consequence of the presence of the size, the threads of the fabric are not able to absorb water, finishing agents or
30 other compositions (e.g. bleaching, dyeing or crease-proofing compositions) to a sufficient degree. Uniform and durable finishing of the fabric can thus be achieved only after removal of the size from the fabric; a process of removing size for this purpose is known as a "desizing" process.

In cases where the size comprises a starch, the desizing treatment may be carried out using a starch-degrading enzyme (e.g. an amylase). In cases where the size comprises fat and/or oil, the desizing treatment may comprise the use of a

5   lipolytic enzyme (a lipase). In cases where the size comprises a significant amount of carboxymethylcellulose (CMC) or other cellulose-derivatives, the desizing treatment may be carried out with a cellulolytic enzyme, either alone or in combination with other substances, optionally in combination with other

10   enzymes, such as amylases and/or lipases.

It is an object of the present invention to achieve improved enzyme performance under desizing conditions by modifying the enzyme so as to alter (increase) the affinity of the enzyme

15   for cellulosic fabric, whereby the modified enzyme comes into closer contact with the sizing agent in question.

SUMMARY OF THE INVENTION

20   It has now surprisingly been found possible to achieve improved enzymatic removal of a sizing agent present on cellulose-containing fabric or textile by means of an enzymatic process wherein the fabric or textile is contacted with an enzyme which has been modified so as to have increased

25   affinity (relative to the unmodified enzyme) for binding to a cellulosic fabric or textile.

3

DETAILED DESCRIPTION OF THE INVENTION

The present invention thus relates, *inter alia*, to a process
for desizing cellulosic fabric or textile, wherein the fabric
5   or textile is treated (normally contacted in aqueous medium)
with a modified enzyme (enzyme hybrid) which comprises a
catalytically (enzymatically) active amino acid sequence of an
enzyme, in particular of a non-cellulolytic enzyme, linked to
an amino acid sequence comprising a cellulose-binding domain.
10

The term "desizing" is intended to be understood in a conven-
tional manner, i.e. the removal of a sizing agent from the
fabric.

15   The terms "cellulose-containing" and "cellulosic" when used
herein in connection with fabric or textile are intended to
indicate any type of fabric, in particular woven fabric,
prepared from a cellulose-containing material, such as cotton,
or from a cellulose-derived material (prepared, e.g., from
20   wood pulp or from cotton).

In the present context, the term "fabric" is intended to
include garments and other types of processed fabrics, and is
used interchangeably with the term "textile".
25

Examples of cellulosic fabric manufactured from naturally
occurring cellulosic fibre are cotton, ramie, jute and flax
(linen) fabrics. Examples of cellulosic fabrics made from man-
made cellulosic fibre are viscose (rayon) and lyocell (e.g.
30   Tencel™) fabric; also of relevance in the context of the
invention are all blends of cellulosic fibres (such as vi-
scose, lyocell, cotton, ramie, jute or flax) with other fi-
bres, such as wool, polyester, polyacrylic, polyamide or po-
lyacetate fibres. Specific examples of blended cellulosic

fabric are viscose/cotton blends, lyocell/cotton blends (e.g. Tencel™/cotton blends), viscose/wool blends, lyocell/wool blends, cotton/wool blends, cotton/polyester blends, viscose/-cotton/polyester blends, wool/cotton/polyester blends, and
5   flax/cotton blends.

Cellulose-binding domains

Although a number of types of carbohydrate-binding domains have been described in the patent and scientific literature,
10 the majority thereof - many of which derive from cellulolytic enzymes (cellulases) - are commonly referred to as "cellulose-binding domains"; a typical cellulose-binding domain (CBD) will thus be one which occurs in a cellulase and which binds preferentially to cellulose and/or to poly- or
15 oligosaccharide fragments thereof.

Cellulose-binding (and other carbohydrate-binding) domains are polypeptide amino acid sequences which occur as integral parts of large polypeptides or proteins consisting of two or
20 more polypeptide amino acid sequence regions, especially in hydrolytic enzymes (hydrolases) which typically comprise a catalytic domain containing the active site for substrate hydrolysis and a carbohydrate-binding domain for binding to the carbohydrate substrate in question. Such enzymes can
25 comprise more than one catalytic domain and one, two or three carbohydrate-binding domains, and they may further comprise one or more polypeptide amino acid sequence regions linking the carbohydrate-binding domain(s) with the catalytic domain(s), a region of the latter type usually being denoted
30 a "linker".

Examples of hydrolytic enzymes comprising a cellulose-binding domain are cellulases, xylanases, mannanases, arabinofuranosidases, acetylesterases and chitinases.

"Cellulose-binding domains" have also been found in algae,
e.g. in the red alga *Porphyra purpurea* in the form of a non-
hydrolytic polysaccharide-binding protein [see P. Tomme et
al., Cellulose-Binding Domains - Classification and
5   Properties, in: Enzymatic Degradation of Insoluble
Carbohydrates, John N. Saddler and Michael H. Penner (Eds.),
ACS Symposium Series, No. 618 (1996)]. However, most of the
known CBDs [which are classified and referred to by P. Tomme
et al. (*op cit.*) as "cellulose-binding domains"] derive from
10   cellulases and xylanases.

In the present context, the term "cellulose-binding domain"
is intended to be understood in the same manner as in the
latter reference (P. Tomme et al., *op. cit*). The P. Tomme et
15   al. reference classifies more than 120 "cellulose-binding
domains" into 10 families (I-X) which may have different
functions or roles in connection with the mechanism of
substrate binding. However, it is to be anticipated that new
family representatives and additional families will appear in
20   the future, and in connection with the present invention a
representative of one such new CBD family has in fact been
identified (see Example 2 herein).

In proteins/polypeptides in which CBDs occur (e.g. enzymes,
25   typically hydrolytic enzymes such as cellulases), a CBD may
be located at the N or C terminus or at an internal position.

That part of a polypeptide or protein (e.g. hydrolytic
enzyme) which constitutes a CBD *per se* typically consists of
30   more than about 30 and less than about 250 amino acid
residues. For example: those CBDs listed and classified in
Family I in accordance with P. Tomme et al. (*op. cit.*)
consist of 33-37 amino acid residues, those listed and
classified in Family IIa consist of 95-108 amino acid

residues, those listed and classified in Family VI consist of
85-92 amino acid residues, whilst one CBD (derived from a
cellulase from *Clostridium thermocellum*) listed and
classified in Family VII consists of 240 amino acid residues.
5   Accordingly, the molecular weight of an amino acid sequence
constituting a CBD *per se* will typically be in the range of
from about 4kD to about 40kD, and usually below about 35kD.

Enzyme hybrids

10 Enzyme classification numbers (EC numbers) referred to in the
present specification with claims are in accordance with the
Recommendations (1992) of the Nomenclature Committee of the
International Union of Biochemistry and Molecular Biology,
Academic Press Inc., 1992.

15

A modified enzyme (enzyme hybrid) for use in accordance with
the invention comprises a catalytically active (enzymatically
active) amino acid sequence (in general a polypeptide amino
acid sequence) of an enzyme, more particularly of a non-
20 cellulolytic enzyme (i.e. a catalytically active amino acid
sequence of an enzyme other than a cellulase), useful in
relation to desizing, in particular of an enzyme selected from
the group consisting of amylases (e.g. α-amylases, EC 3.2.1.1)
and lipases (e.g. triacylglycerol lipases, EC 3.1.1.3), fused
25 (linked) to an amino acid sequence comprising a cellulose-
binding domain. The catalytically active amino acid sequence
in question may comprise or consist of, for example, the whole
of - or substantially the whole of - the full amino acid
sequence of the mature enzyme in question, or it may consist
30 of a portion of the full sequence which retains substantially
the same catalytic (enzymatic) properties as the full
sequence.

Modified enzymes (enzyme hybrids) of the type in question, as

well as detailed descriptions of the preparation and purification thereof, are known in the art [see, e.g., WO 90/00609, WO 94/24158 and WO 95/16782, as well as Greenwood et al., _Biotechnology and Bioengineering_ 44 (1994) pp. 1295-

5   1305]. They may, e.g., be prepared by transforming into a host cell a DNA construct comprising at least a fragment of DNA encoding the cellulose-binding domain ligated, with or without a linker, to a DNA sequence encoding the enzyme of interest, and growing the transformed host cell to express the fused

10   gene. One relevant, but non-limiting, type of recombinant product (enzyme hybrid) obtainable in this manner - often referred to in the art as a "fusion protein" - may be described by one of the following general formulae:

15   A-CBD-MR-X-B

                A-X-MR-CBD-B

      In the latter formulae, CBD is an amino acid sequence
20   comprising at least the cellulose-binding domain (CBD) _per se_.

      MR (the middle region; a linker) may be a bond, or a linking group comprising from 1 to about 100 amino acid residues, in particular of from 2 to 40 amino acid residues, e.g. from 2 to

25   15 amino acid residues. MR may, in principle, alternatively be a non-amino-acid linker.

      X is an amino acid sequence comprising the above-mentioned, catalytically (enzymatically) active sequence of amino acid

30   residues of a polypeptide encoded by a DNA sequence encoding the non-cellulolytic enzyme of interest.

      The moieties A and B are independently optional. When present, a moiety A or B constitutes a terminal extension of a CBD or X

35   moiety, and normally comprises one or more amino acid

8

residues.

It will thus, *inter alia*, be apparent from the above that a
CBD in an enzyme hybrid of the type in question may be
5   positioned C-terminally, N-terminally or internally in the
enzyme hybrid. Correspondingly, an X moiety in an enzyme
hybrid of the type in question may be positioned N-terminally,
C-terminally or internally in the enzyme hybrid.

10 Enzyme hybrids of interest in the context of the invention
include enzyme hybrids which comprise more than one CBD, e.g.
such that two or more CBDs are linked directly to each other,
or are separated from one another by means of spacer or linker
sequences (consisting typically of a sequence of amino acid
15 residues of appropriate length). Two CBDs in an enzyme hybrid
of the type in question may, for example, also be separated
from one another by means of an -MR-X- moiety as defined
above.

20 A very important issue in the construction of enzyme hybrids
of the type in question is the stability towards proteolytic
degradation. Two- and multi-domain proteins are particularly
susceptible towards proteolytic cleavage of linker regions
connecting the domains. Proteases causing such cleavage may,
25 for example, be subtilisins, which are known to often exhibit
broad substrate specificities [see, e.g.: Grøn et al.,
Biochemistry 31 (1992), pp. 6011-6018; Teplyakov et al.,
Protein Engineering 5 (1992), pp. 413-420].

30 Glycosylation of linker residues in eukaryotes is one of
Nature's ways of preventing proteolytic degradation. Another
is to employ amino acids which are less favoured by the
surrounding proteases. The length of the linker also plays a
role in relation to accessibility by proteases. Which
35 "solution" is optimal depends on the environment in which the

9

enzyme hybrid is to function.

When constructing new enzyme hybrid molecules, linker stability
5   thus becomes an issue of great importance. The various linkers described in examples presented herein (vide infra) in the context of the present invention are intended to take account of this issue.

10   <u>Cellulases (cellulase genes) useful for preparation of CBDs</u>
Techniques suitable for isolating a cellulase gene are well known in the art. In the present context, the terms "cellulase" and "cellulolytic enzyme" refer to an enzyme which catalyses the degradation of cellulose to glucose, cellobiose,
15   triose and/or other cello-oligosaccharides.

Preferred cellulases (i.e. cellulases comprising preferred CBDs) in the present context are microbial cellulases, particularly bacterial or fungal cellulases. Endoglucanases,
20   notably endo-1,4-β-glucanases (EC 3.2.1.4), particularly monocomponent (recombinant) endo-1,4-β-glucanases, are a preferred class of cellulases,.

Useful examples of bacterial cellulases are cellulases derived
25   from or producible by bacteria from the group consisting of Pseudomonas, Bacillus, Cellulomonas, Clostridium, Microspora, Thermotoga, Caldocellum and Actinomycets such as Streptomyces, Termomonospora and Acidothemus, in particular from the group consisting of Pseudomonas cellulolyticus, Bacillus lautus,
30   Cellulomonas fimi, Clostridium thermocellum, Microspora bispora, Termomonospora fusca, Termomonospora cellulolyticum and Acidothemus cellulolyticus.

10

The cellulase may be an acid, a neutral or an alkaline cellulase, i.e. exhibiting maximum cellulolytic activity in the acid, neutral or alkaline range, respectively.

5   A useful cellulase is an acid cellulase, preferably a fungal acid cellulase, which is derived from or producible by fungi from the group of genera consisting of *Trichoderma*, *Myrothecium*, *Aspergillus*, *Phanaerochaete*, *Neurospora*, *Neocallimastix* and *Botrytis*.

10

A preferred useful acid cellulase is one derived from or producible by fungi from the group of species consisting of *Trichoderma viride*, *Trichoderma reesei*, *Trichoderma longibra-chiatum*, *Myrothecium verrucaria*, *Aspergillus niger*, *Aspergil-*
15   *lus oryzae*, *Phanaerochaete chrysosporium*, *Neurospora crassa*, *Neocallimastix partriciarum* and *Botrytis cinerea*.

Another useful cellulase is a neutral or alkaline cellulase, preferably a fungal neutral or alkaline cellulase, which is
20   derived from or producible by fungi from the group of genera consisting of *Aspergillus*, *Penicillium*, *Myceliophthora*, *Humicola*, *Irpex*, *Fusarium*, *Stachybotrys*, *Scopulariopsis*, *Chaetomium*, *Mycogone*, *Verticillium*, *Myrothecium*, *Papulospora*, *Gliocladium*, *Cephalosporium* and *Acremonium*.
25

A preferred alkaline cellulase is one derived from or produ-cible by fungi from the group of species consisting of *Humicola insolens*, *Fusarium oxysporum*, *Myceliopthora ther-mophila*, *Penicillium janthinellum* and *Cephalosporium* sp.,
30   preferably from the group of species consisting of *Humicola insolens* DSM 1800, *Fusarium oxysporum* DSM 2672, *Myceliopthora thermophila* CBS 117.65, and *Cephalosporium* sp. RYM-202.

A preferred cellulase is an alkaline endoglucanase which is

immunologically reactive with an antibody raised against a highly purified ~43kD endoglucanase derived from *Humicola insolens* DSM 1800, or which is a derivative of the latter ~43kD endoglucanase and exhibits cellulase activity.

5

Other examples of useful cellulases are variants of parent cellulases of fungal or bacterial origin, e.g. variants of a parent cellulase derivable from a strain of a species within the fungal genera *Humicola*, *Trichoderma*, *Fusarium* or

10 *Myceliophthora*.

Isolation of a cellulose-binding domain

In order to isolate a cellulose-binding domain of, e.g., a cellulase, several genetic engineering approaches may be used.

15 One method uses restriction enzymes to remove a portion of the gene and then to fuse the remaining gene-vector fragment in frame to obtain a mutated gene that encodes a protein truncated for a particular gene fragment. Another method involves the use of exonucleases such as *Bal31* to

20 systematically delete nucleotides either externally from the 5' and the 3' ends of the DNA or internally from a restricted gap within the gene. These gene-deletion methods result in a mutated gene encoding a shortened gene molecule whose expression product may then be evaluated for substrate-binding

25 (e.g. cellulose-binding) ability. Appropriate substrates for evaluating the binding ability include cellulosic materials such as Avicel™ and cotton fibres. Other methods include the use of a selective or specific protease capable of cleaving a CBD, e.g. a terminal CBD, from the remainder of the

30 polypeptide chain of the protein in question.

As already indicated (*vide supra*), once a nucleotide sequence encoding the substrate-binding (carbohydrate-binding) region has been identified, either as cDNA or chromosomal DNA, it may

then be manipulated in a variety of ways to fuse it to a DNA
sequence encoding the enzyme or enzymatically active amino
acid sequence of interest. The DNA fragment encoding the
carbohydrate-binding amino acid sequence, and the DNA encoding
5   the enzyme or enzymatically active amino acid sequence of
interest are then ligated with or without a linker. The
resulting ligated DNA may then be manipulated in a variety of
ways to achieve expression. Preferred microbial expression
hosts include certain *Aspergillus* species (e.g. *A. niger* or *A.*
10   *oryzae*), *Bacillus* species, and organisms such as *Escherichia*
coli or *Saccharomyces cerevisiae*.

Amylolytic enzymes

Amylases (e.g. α- or β-amylases) which are appropriate as the
15   basis for enzyme hybrids of the types employed in the context
of the present invention include those of bacterial or fungal
origin. Chemically or genetically modified mutants of such
amylases are included in this connection. Relevant α-amylases
include, for example, α-amylases obtainable from *Bacillus*
20   species, in particular a special strain of *B. licheniformis*,
described in more detail in GB 1296839. Relevant commercially
available amylases include Duramyl™, Termamyl™, Fungamyl™ and
BAN™ (all available from Novo Nordisk A/S, Bagsvaerd,
Denmark), and Rapidase™ and Maxamyl™ P (available from Gist-
25   Brocades, Holland).

Other useful amylolytic enzymes are CGTases (cyclodextrin
glucanotransferases, EC 2.4.1.19), e.g. those obtainable from
species of *Bacillus*, *Thermoanaerobactor* or *Thermoanaero-*
30   *bacterium*.

Lipolytic enzymes

Lipolytic enzymes (lipases) which are appropriate as the basis

for enzyme hybrids of the types employed in the context of the
present invention include those of bacterial or fungal origin.
Chemically or genetically modified mutants of such lipases are
included in this connection.

5

Examples of useful lipases include a *Humicola lanuginosa*
lipase, e.g. as described in EP 258 068 and EP 305 216; a
Rhizomucor miehei lipase, e.g. as described in EP 238 023; a
Candida lipase, such as a *C. antarctica* lipase, e.g. the *C.*

10 *antarctica* lipase A or B described in EP 214 761; a
Pseudomonas lipase, such as one of those described in EP 721
981 (e.g. a lipase obtainable from a *Pseudomonas* sp. SD705
strain having deposit accession number FERM BP-4772), in
PCT/JP96/00426, in PCT/JP96/00454 (e.g. a *P. solanacearum*

15 lipase), in EP 571 982 or in WO 95/14783 (e.g. a *P. mendocina*
lipase), a *P. alcaligenes* or *P. pseudoalcaligenes* lipase, e.g.
as described in EP 218 272, a *P. cepacia* lipase, e.g. as
described in EP 331 376, a *P. stutzeri* lipase, e.g. as
disclosed in GB 1,372,034, or a *P. fluorescens* lipase; a

20 *Bacillus* lipase, e.g. a *B. subtilis* lipase [Dartois et al.,
Biochemica et Biophysica Acta 1131 (1993) pp. 253-260], a *B.*
stearothermophilus lipase (JP 64/744992) and a *B. pumilus*
lipase (WO 91/16422).

25 Furthermore, a number of cloned lipases may be useful, includ-
ing the *Penicillium camembertii* lipase described by Yamaguchi
et al. in Gene 103 (1991), pp. 61-67, the *Geotricum candidum*
lipase [Y. Schimada et al., J. Biochem. 106 (1989), pp. 383-
388], and various *Rhizopus* lipases such as an *R. delemar*

30 lipase [M.J. Hass et al., Gene 109 (1991) pp. 117-113], an *R.*
niveus lipase [Kugimiya et al., Biosci. Biotech. Biochem. 56
(1992), pp. 716-719] and a *R. oryzae* lipase.

14

Other potentially useful types of lipolytic enzymes include cutinases, e.g. a cutinase derived from *Pseudomonas mendocina* as described in WO 88/09367, or a cutinase derived from
5 *Fusarium solani* f. pisi (described, e.g., in WO 90/09446).

Suitable commercially available lipases include Lipolase™ and Lipolase Ultra™ (available from Novo Nordisk A/S), M1 Lipase™, Lumafast™ and Lipomax™ (available from Gist-Broca-
10 des) and Lipase P "Amano" (available from Amano Pharmaceutical Co. Ltd.).

Plasmids

Preparation of plasmids capable of expressing fusion proteins
15 having the amino acid sequences derived from fragments of more than one polypeptide is well known in the art (see, for example, WO 90/00609 and WO 95/16782). The expression cassette may be included within a replication system for episomal maintenance in an appropriate cellular host or may be provided
20 without a replication system, where it may become integrated into the host genome. The DNA may be introduced into the host in accordance with known techniques such as transformation, microinjection or the like.

25 Once the fused gene has been introduced into the appropriate host, the host may be grown to express the fused gene. Normally it is desirable additionally to add a signal sequence which provides for secretion of the fused gene. Typical examples of useful fused genes are:
30

Signal   sequence   --   (pro-peptide)   --   carbohydrate-binding domain -- linker -- enzyme sequence of interest, or

Signal   sequence   --   (pro-peptide)   --   enzyme   sequence   of

interest -- linker -- carbohydrate-binding domain,

in which the pro-peptide sequence normally contains 5-100, e.g. 5-25, amino acid residues.

The recombinant product may be glycosylated or non-glycosylated.

Determination of α-amylolytic activity (KNU)

The α-amylolytic activity of an enzyme or enzyme hybrid may be determined using potato starch as substrate. This method is based on the break-down (hydrolysis) of modified potato starch, and the reaction is followed by mixing samples of the starch/enzyme or starch/enzyme hybrid solution with an iodine solution. Initially, a blackish-blue colour is formed, but during the break-down of the starch the blue colour becomes weaker and gradually turns to a reddish-brown. The resulting colour is compared with coloured glass calibration standards.

One Kilo Novo α-Amylase Unit (KNU) is defined as the amount of enzyme (enzyme hybrid) which, under standard conditions (i.e. at $37\pm0.05°C$, 0.0003 M Ca^{2+}, pH 5.6) dextrinizes 5.26 g starch dry substance (Merck Amylum solubile) per hour.

Determination of lipolytic activity (LU)
The lipolytic (lipase) activity of an enzyme or enzyme hybrid may be determined using tributyrin (glyceryl tributyrate) as substrate. This method is based on the hydrolysis of tributyrin by the enzyme or enzyme hybrid, and the alkali consumption is registered as a function of time.

One Lipase Unit (LU) is defined as the amount of enzyme (enzyme hybrid) which, under standard conditions (i.e. at 30.0°C, pH 7.0; with Gum Arabic as emulsifier and tributyrin

as substrate) liberates 1 μmol of titratable butyric acid per minute.

Process conditions

5   It will be understood that the method of the invention may be performed in accordance with any suitable desizing procedure known in the art, e.g. as described by E.S. Olson in Textile Wet Processes, Vol. I, Noyes Publication, Park Ridge, New Jersey, USA (1983), or by M. Peter and H.K. Rouette in
10   Grundlagen der Textilveredlung, Deutsche Fachverlag GmbH, Frankfurt am Main, Germany (1988). Thus, the process conditions to be used in performing the present invention may be selected so as to match particular equipment or a particular type of process which it is desirable to use. Preferred
15   examples of types of procedures suitable for use in connection with the present invention include Jigger/Winch, Pad-Roll and Pad-Steam types. These types are dealt with in further detail below.

20   The process of the invention may be carried out as a batch, semi-continuous or continuous process. As an example, the process may comprise the following steps:

     (a) impregnating the fabric in a desizing bath containing (as
25   a minimum) an amylolytic enzyme hybrid and/or a lipolytic enzyme hybrid;

     (b) subjecting the impregnated fabric to steaming, so as to bring the fabric to the desired reaction temperature, gen-
30   erally between 20° and 120°C; and

     (c) holding by rolling-up or pleating the cloth in a J-Box, U-Box, carpet machine or the like for a sufficient period of time (normally between a few minutes and several hours) to

17

allow the desizing to occur.

Prior to carrying out the chosen treatment, the amylolytic enzyme hybrid and/or the lipolytic enzyme hybrid may
5   conveniently be mixed with other components which are conventionally used in the desizing process.

Further components required for performance of the process may be added separately. Thus, for example, a wetting agent and,
10 optionally, a stabilizer may be added. The stabilizer in question may be an agent stabilizing the amylolytic enzyme hybrid and/or the lipolytic enzyme hybrid. Wetting agents serve to improve the wettability of the fibre, whereby rapid and even desizing may be achieved. The wetting agent is
15 preferably of an oxidation-stable type.

In a preferred embodiment of the process of the invention, an amylolytic enzyme hybrid is used in an amount corresponding to an amylase activity in the range of between 1 and 5000 KNU per
20 litre of desizing liquor, such as between 10 and 1000 KNU per litre of desizing liquor, preferably between 50 and 500 KNU per litre, more preferably between 20 and 500 KNU per litre of desizing liquor.

25 In a preferred embodiment of the process of the invention, a lipolytic enzyme hybrid is used in an amount corresponding to a lipase activity in the range of between 10 and 20000 LU per litre of desizing liquor, such as between 50 and 10000 LU per litre of desizing liquor, more preferably between 100 and 5000
30 LU per litre of desizing liquor.

Irrespective of the particular type of procedure to be used for the desizing, the process of the invention is normally performed at a temperature in the range of 30-100°C, such as

18

35-60°C, and at a pH in the range of 3-11, preferably 7-9. However, the actual process conditions may vary widely within these ranges.

5   It will be understood that the process may be performed in any equipment sufficiently tolerant towards the process conditions in question.

The process of the invention may be employed alone or in
10   combination with one or more other enzymatic desizing processes. Suitable combinations include the following:

a treatment with an amylolytic enzyme hybrid, and a treatment with a cellulase;
15   a treatment with a lipolytic enzyme hybrid, and a treatment with a cellulase;
a treatment with an amylolytic enzyme hybrid, and a treatment with a lipase or a lipolytic enzyme hybrid;
a treatment with a lipolytic enzyme hybrid, and a treatment
20   with an amylase or an amylolytic enzyme hybrid;
a treatment with an amylolytic enzyme hybrid, and a treatment with a lipase or a lipolytic enzyme hybrid, and a treatment with a cellulase;
a treatment with a lipolytic enzyme hybrid, and a treatment
25   with an amylase or an amylolytic enzyme hybrid, and a treatment with a cellulase.

The various enzymes/enzyme hybrids will normally be added in one step, but the desizing process may also be performed in
30   more than one step, taking one enzyme/enzyme hybrid at a time.

Composition of the invention
Although an enzyme hybrid, e.g. amylolytic enzyme hybrid and/or lipolytic enzyme hybrid, may be added as such, it is

preferred that it is formulated in the form of a suitable desizing composition.

The desizing composition of the invention may comprise a
5 single type of enzyme hybrid, or more than one type of enzyme hybrid (e.g. an amylolytic enzyme hybrid together with a lipolytic enzyme hybrid). The composition may be in the form of, e.g., a granulate, preferably a non-dusting granulate, or a liquid, in particular a stabilized liquid, or a slurry, or
10 in a protected form. Non-dusting granulates may be produced, for example, as disclosed in US 4,106,991 and US 4,661,452 (both to Novo Nordisk A/S) and may optionally be coated by methods known in the art. In the case of granular formulations ("granulates"), different enzyme hybrids may be formulated,
15 for example, either as a single granulate wherein the individual granules each contain all the enzyme hybrids in question, or as a mixture of discrete, different granulates wherein the individual granules each contain one type of enzyme hybrid of the kind in question.
20

Liquid enzyme preparations may, for instance, be stabilized by adding a polyol (such as propylene glycol or another glycol), a sugar, a sugar alcohol or acetic acid, according to established procedures. Other enzyme stabilizers are well
25 known in the art. Protected enzymes may be prepared as disclosed in EP 238 216.

The composition of the invention may comprise a wetting agent and/or, optionally, one or more further components selected
30 from the group consisting of dispersing agents, sequestering agents (and/or precipitants) and emulsifying agents. An example of a suitable wetting agent is the commercial product Arbyl™ R, available from Grünau, Germany. An emulsifying agent serves to emulsify hydrophobic impurities which may be

20

present on the fabric. A dispersing agent serves to prevent the redeposition of extracted impurities on the fabric. A sequestering agent or precipitant serves to remove metal ions (such as Ca^{2+}, Mg^{2+} and Fe^{2+}) which may have a negative impact on

5   the process; suitable examples include caustic soda (sodium hydroxide) and soda ash (sodium carbonate).

A further aspect of the invention relates to a DNA construct disclosed herein which encodes, or which comprises a sequence

10   which encodes, an enzyme hybrid as disclosed in the present specification.

A still further aspect of the invention relates to a polypeptide (fusion protein or enzyme hybrid) which is encoded

15   by such a DNA construct or sequence, and/or which is disclosed in the present specification.

The invention is further illustrated by means of the examples given below, which are in no way intended ': limit the scope

20   of the invention as claimed:

MATERIALS AND METHODS

Strains:

25   *Bacillus agaradherens* NCIMB No. 40482: comprises the endoglucanase enzyme encoding DNA sequence of Example 2, below.

Escherichia coli SJ2 [Diderichsen et al., J. Bacteriol. 172

30   (1990), pp. 4315-4321].

Electrocompetent cells prepared and transformed using a Bio-Rad GenePulser™ as recommended by the manufacturer.

21

Bacillus subtilis PL2306: this strain is the *B.subtilis* DN1885 with disrupted apr and npr genes [Diderichsen et al., J. Bacteriol. 172 (1990), pp. 4315-4321] disrupted in the transcriptional unit of the known *Bacillus subtilis* cellulase

5   gene, resulting in cellulase-negative cells. The disruption was performed essentially as described in Sonenshein et al. (Eds.), *Bacillus subtilis and other Gram-Positive Bacteria*, American Society for Microbiology (1993), p.618.

10 **Plasmids:**

pDN1528 [Jørgensen et al., J. Bacteriol. 173 (1991), p.559-567].

pBluescriptKSII- (Stratagene, USA).
15

pDN1981 [Jørgensen et al., Gene 96 (1990), p37-41].

Solutions/Media

TY and LB agar [as described in Ausubel et al. (Eds.),
20 *Current Protocols in Molecular Biology*, John Wiley and Sons (1995)].

SB: 32 g Tryptone, 20 g yeast extract, 5 g sodium chloride and 5 ml 1 N sodium hydroxide are mixed in sterile water to a
25 final volume of 1 litre. The solution is sterilised by autoclaving for 20 minutes at 121°C.

10% Avicel™: 100 g of Avicel™ (FLUKA, Switzerland) is mixed with sterile water to a final volume of 1 litre, and the
30 resulting 10% Avicel™ is sterilised by autoclaving for 20 minutes at 121°C.

Buffer: 0.05 M potassium phosphate, pH 7.5.

22

General molecular biology methods

DNA manipulations and transformations were performed using
standard methods of molecular biology [Sambrook et al.,
Molecular Cloning: A Laboratory Manual, Cold Spring Harbor
5   lab., Cold Spring Harbor, NY (1989); Ausubel et al. (Eds.),
Current Protocols in Molecular Biology, John Wiley and Sons
(1995); C.R. Harwood and S.M. Cutting (Eds.) Molecular
Biological Methods for Bacillus, John Wiley and Sons (1990)].

10   Enzymes for DNA manipulations were used according to the
specifications of the suppliers.

EXAMPLE 1

15   **Subcloning of a partial Termamyl sequence.**
The alfa-amylase gene encoded on pDN1528 was PCR amplified
for introduction of a BamHI site in the 3'-end of the coding
region. The PCR and the cloning were carried out as follows:

20   Approximately 10-20 ng of plasmid pDN1528 was PCR amplified
in HiFidelity™ PCR buffer (Boehringer Mannheim, Germany)
supplemented with 200 µM of each dNTP, 2.6 units of
HiFidelity™ Expand enzyme mix, and 300 pmol of each primer:

25   #5289
5'-GCT TTA CGC CCG ATT GCT GAC GCT G -3'

30   #26748
5'-GCG ATG AGA CGC GCG GCC GCC TAT CTT TGA ACA TAA ATT GAA
ACG GAT CCG -3'

(BamHI restriction site underlined].

23

The PCR reactions were performed using a DNA thermal cycler
(Landgraf, Germany). One incubation at 94°C for 2 min, 60°C
for 30 sec and 72°C for 45 sec was followed by ten cycles of
5   PCR performed using a cycle profile of denaturation at 94°C
for 30 sec, annealing at 60°C for 30 sec, and extension at
72°C for 45sec and twenty cycles of denaturation at 94°C for
30 sec, 60°C for 30 sec and 72°C for 45 sec (at this
elongation step, 20 sec are added every cycle). 10 µl
10   aliquots of amplification product were analyzed by
electrophoresis in 1.0 % agarose gels (NuSieve™, FMC) with
ReadyLoad™ 100bp DNA ladder (GibcoBRL, Denmark) as a size
marker.

15   40 µl aliquots of PCR product generated as described above
were purified using QIAquick™ PCR purification kit (Qiagen,
USA) according to the manufacturer's instructions. The
purified DNA was eluted in 50 µl of 10mM Tris-HCl, pH 8.5. 25
µl of the purified PCR fragment was digested with BamHI and
20   PstI, subjected to electrophoresis in 1.0% low gelling
temperature agarose (SeaPlaque™ GTG, FMC) gels, and the
relevant fragment was excised from the gel and purified using
QIAquick™ Gel extraction Kit (Qiagen, USA) according to the
manufacturer's instructions. The isolated DNA fragment was
25   then ligated to BamHI-PstI digested pBluescriptII KS-, and
the ligation mixture was used to transform E. coli SJ2.

Cells were plated on LB agar plates containing Ampicillin
(200 µg/ml) and supplemented with X-gal (5-bromo-4-chloro-3-
30   indolyl-α-D-galactopyranoside, 50 µg/ml), and incubated at
37°C overnight. The next day, white colonies were restreaked
onto fresh LB-Ampicillin agar plates and incubated at 37°C

24

overnight. The following day, single colonies were
transferred to liquid LB medium containing Ampicillin (200
µg/ml) and incubated overnight at 37°C with shaking at 250
rpm.

5

Plasmids were extracted from the liquid cultures using QIAgen
Plasmid Purification mini kit (Qiagen, USA) according to the
manufacturer's instructions. 5 µl samples of the plasmids
were digested with PstI and BamHI. The digestions were
10   checked by gel electrophoresis on a 1.0% agarose gel (NuSie-
ve™, FMC). One positive clone, containing the PstI-BamHI
fragment containing part of the α-amylase gene, was
designated pMB335. This plasmid was then used in the
construction of α-amylase-CBD hybrid.

15

Isolation of genomic DNA
Clostridium stercorarium NCIMB 11754 was grown anaerobically
at 60°C in specified media as recommended by The National
Collections of Industrial and Marine Bacteria Ltd. (NCIMB),
20   Scotland. Cells were harvested by centrifugation.

Genomic DNA was isolated as described by Pitcher et al, Lett.
Appl. Microbiol. 8 (1989), pp. 151-156.

25   **In vitro amplification of the CBD-dimer of *Clostridium***
stercorarium (NCIMB 11754) XynA
Approximately 100-200 ng of genomic DNA was PCR amplified in
HiFidelity™ PCR buffer (Boehringer Mannheim, Germany)
supplemented with 200 µM of each dNTP, 2.6 units of
30   HiFidelity™ Expand enzyme mix, and 300 pmol of each primer:

#27183
5'-GCT GCA GGA TCC GTT TCA ATT TAT GTT CAA AGA TCT GGC GGA

CCT GGA ACG CCA AAT AAT GGA AGA GG -3'

#27182
5'-GCA CTA GCT AGA CGG CCG CTA CCA GTC AAC ATT AAC AGG ACC
5   TGA G -3'

(BamHI and EagI restriction sites underlined).

The primers were designed to amplify the DNA encoding the
10   cellulose-binding domain of the XynA-encoding gene of
     Clostridium stercorarium NCIMB 11754; the DNA sequence was
     extracted from the database GenBank under the accession
     number D13325.

15   The PCR reactions were performed using a DNA thermal cycler
     (Landgraf, Germany). One incubation at 94°C for 2 min, 60°C
     for 30 sec and 72°C for 45 sec was followed by ten cycles of
     PCR performed using a cycle profile of denaturation at 94°C
     for 30 sec, annealing at 60°C for 30 sec, and extension at
20   72°C for 45 sec and twenty cycles of denaturation at 94°C for
     30 sec, 60°C for 30 sec and 72°C for 45 sec (at this
     elongation step, 20 sec are added every cycle). 10 µl
     aliquots of amplification product were analyzed by
     electrophoresis in 1.0 % agarose gels (NuSieve™, FMC) with
25   ReadyLoad™ 100bp DNA ladder (GibcoBRL, Denmark) as a size
     marker.

Cloning by polymerase chain reaction (PCR):
Subcloning of PCR fragments.
30   40 µl aliquots of PCR product generated as described above
     were purified using QIAquick™ PCR purification kit (Qiagen,
     USA) according to the manufacturer's instructions. The
     purified DNA was eluted in 50 µl of 10mM Tris-HCl, pH 8.5. 25

26

μl of the purified PCR fragment was digested with BamHI and
EagI, subjected to electrophoresis in 1.0% low gelling
temperature agarose (SeaPlaque™ GTG, FMC) gels, and the
relevant fragment was excised from the gels and purified
5   using QIAquick™ Gel extraction Kit (Qiagen, USA) according
to the manufacturer's instructions. The isolated DNA fragment
was then ligated to BamHI-NotI digested pMB335 and the
ligation mixture was used to transform E.coli SJ2.

10   **Identification and characterization of positive clones**
Cells were plated on LB agar plates containing Ampicillin
(200 μg/ml) and incubated at 37°C overnight. The next day,
colonies were restreaked onto fresh LB-Ampicillin agar plates
and incubated at 37°C overnight. The following day, single
15   colonies were transferred to liquid LB medium containing
Ampicillin (200 μg/ml) and incubated overnight at 37°C with
shaking at 250 rpm.

Plasmids were extracted from the liquid cultures using QIAgen
20   Plasmid Purification mini kit (Qiagen, USA) according to the
manufacturer's instructions. 5 μl samples of the plasmids
were digested with BamHI and NotII. The digestions were
checked by gel electrophoresis on a 1.0% agarose gel (NuSie-
ve™, FMC). The appearance of a DNA fragment of the same
25   size as seen from the PCR amplification indicated a positive
clone.

One positive clone, containing the fusion construct of the α-
amylase gene and the CBD-dimer of Clostridium stercorarium
30   (NCIMB 11754) XynA, was designated MBamyX.

**Cloning of the fusion construct into a Bacillus-based
expression vector**

27

The pDN1528 vector contains the amyL gene of *B.licheniformis*; this gene is actively expressed in *B. subtilis*, resulting in the production of active α-amylase appearing in the supernatant. For expression purposes, the DNA encoding the
5   fusion protein as constructed above was introduced to pDN1528.
This was done by digesting pMBamyX and pDN1528 with SalI-NotI, purifying the fragments and ligating the 4.7 kb pDN1528 SalI-NotI fragment with the 1.0 kb pMBamyX SalI-NotI
10   fragment. This created an inframe fusion of the hybrid construction with the Termamyl™ (*B. licheniformis* α-amylase) gene. The DNA sequence of the fusion construction of pMB206, and the corresponding amino acid sequence, are shown in SEQ ID No. 1 and SEQ ID No. 2, respectively.
15

The ligation mixture was used to transform competent cells of *B. subtilis* PL2306. Cells were plated on LB agar plates containing chloramphenicol (6 µg/ml), 0.4% glucose and 10mM potassium hydrogen phosphate, and incubated at 37°C
20   overnight. The next day, colonies were restreaked onto fresh LBPG (LB plates with 0.4% glucose and 10mM potassium phosphate, pH 10) chloramphenicol agar plates and incubated at 37°C overnight. The following day, single colonies of each clone were transferred to liquid LB medium containing
25   chloramphenicol (6 µg/ml) and incubated overnight at 37°C with shaking at 250 rpm.

Plasmids were extracted from the liquid cultures using QIAgen Plasmid Purification mini kit (Qiagen, USA) according to the
30   manufacturer's instructions. However, the resuspension buffer was supplemented with 1 mg/ml of chicken egg white lysozyme (SIGMA, USA) prior to lysing the cells at 37°C for 15 minutes. 5 µl samples of the plasmids were digested with

28

BamHI and NotI. The digestions were checked by gel
electrophoresis on a 1.5% agarose gel (NuSieve™, FMC). The
appearance of a DNA fragment of the same size as seen from
the PCR amplification indicated a positive clone. One
5   positive clone was designated MB-BSamyx.

**Expression, secretion and functional analysis of the fusion
protein**

The clone MB-BSamyx (expressing Termamyl™ fused to
10   C.stercorarium XynA dimer CBD) was incubated for 20 hours in
SB medium at 37°C with shaking at 250 rpm. 1 ml of cell-free
supernatant was mixed with 200 μl of 10% Avicel™. The
mixture was incubated for 1 hour at 0°C and then centrifuged
for 5 minutes at 5000 x g. The pellet was resuspended in 100
15   μl of SDS-PAGE buffer, and the suspension was boiled at 95°C
for 5 minutes, centrifuged at 5000 x g for 5 minutes, and 25
μl was loaded onto a 4-20% Laemmli Tris-Glycine, SDS-PAGE
NOVEX™ gel (Novex, USA). The samples were subjected to
electrophoresis in an Xcell™ Mini-Cell (NOVEX, USA) as
20   recommended by the manufacturer. All subsequent handling of
gels, including staining (Coomassie), destaining and drying,
were performed as described by the manufacturer.

The appearance of a protein band of molecular weight approx.
25   85 kDa indicated expression in B.subtilis of the Termamyl-CBD
fusion amyx.

EXAMPLE 2

30   **Identification of a novel CBD representing a new CBD family**
The alkaline cellulase cloned in Bacillus subtilis as
described below was expressed by incubating the clone for 20
hours in SB medium at 37°C with shaking at 250 rpm. The

expressed cellulase was shown to contain a CBD by its ability
to specifically bind to Avicel™.

5   When left to incubate for a further 20 hours, the cellulase
was proteolytically cleaved and two specific protein bands
appeared in SDS-PAGE, one corresponding to the catalytic part
of the cellulase, approximate molecular weight (MW) 35 kD,
and the other corresponding to a proposed linker and CBD of
approximate MW 8 kD.

10

The CBD was found to be the C-terminal part of the cellulase,
and did not match any of the CBD families described
previously [Tomme et al., Cellulose-Binding Domains:
Classification and Properties, In: J.N. Saddler and M.H.

15   Penner (Eds.), Enzymatic Degradation of Insoluble
Carbohydrates, ACS Symposium Series No. 618 (1996)].
Accordingly, this CBD appears to be the first member of a new
family.

20   **Cloning of the alkaline cellulase (endoglucanase) from
Bacillus agaradherens and expression of the alkaline
cellulase in Bacillus subtilis**
The nucleotide sequence encoding the alkaline cellulase from
Bacillus agaradherens (deposited under accession No. NCIMB

25   40482) was cloned by PCR for introduction in an expression
plasmid pDN1981. PCR was performed essentially as described
above on 500 ng of genomic DNA, using the following two
primers containing NdeI and KpnI restriction sites for
introducing the endoglucanase-encoding DNA sequence to

30   pDN1981 for expression:

#20887
5'-GTA GGC TCA GTC ATA TGT TAC ACA TTG AAA GGG GAG GAG AAT
CAT GAA AAA GAT AAC TAC TAT TTT TGT CG-3'

#21318
5'-GTA CCT CGC G<u>GG TAC C</u>AA GCG GCC GCT TAA TTG AGT GGT TCC
CAC GGA CCG-3'

5

After PCR cycling, the PCR fragment was purified using QIA-
quick™ PCR column kit (Qiagen, USA) according to the
manufacturer's instructions. The purified DNA was eluted in 50
µl of 10mM Tris-HCl, pH 8.5, digested with NdeI and KpnI,
10   purified and ligated to digested pDN1981. The ligation mixture
was used to transform *B. subtilis* PL2306. Competent cells were
prepared and transformed as described by Yasbin et al., <u>J.
Bacteriol.</u> <u>121</u> (1975), pp. 296-304.

15   **Isolation and testing of *B. subtilis* transformants**
The transformed cells were plated on LB agar plates containing
Kanamycin (10 mg/ml), 0.4% glucose, 10 mM potassium phosphate
and 0.1% AZCL HE-cellulose (Megazyme, Australia), and
incubated at
20   37 °C for 18 hours. Endoglucanase-positive colonies were
identified as colonies surrounded by a blue halo.

Each of the positive transformants was inoculated in 10 ml TY
medium containing Kanamycin (10 mg/ml). After 1 day of incuba-
25   tion at 37°C with shaking at 250rpm, 50 ml of supernatant was
removed. The endoglucanase activity was identified by adding
50 ml of supernatant to holes punctured in the agar of LB agar
plates containing 0.1% AZCL HE-cellulose.

30   After 16 hours incubation at 37°C, blue halos surrounding
holes indicated expression of the endoglucanase in *B.
subtilis*. One such clone was designated MB208. The encoding
DNA sequence and amino acid sequence of the endoglucanase are
shown in SEQ ID No. 3 and SEQ ID No. 4, respectively.

The DNA sequence was determined as follows: Qiagen purified
plasmid DNA was sequenced with the Taq deoxy terminal cycle
sequencing kit (Perkin Elmer, USA) using the primers #21318
5   and #20887 (*vide supra*) and employing an Applied Biosystems
373A automated sequencer operated according to the manufac-
turer's instructions. Analysis of the sequence data is
performed according to Devereux et al., <u>Carcinogenesis</u> <u>14</u>
(1993), pp. 795-801.

10

**In vitro amplification of the CBD of *Bacillus agaradherens*
NCIMB 40482 endoglucanase**
Approximately 10-20 ng of plasmid pMB208 was PCR amplified in
HiFidelity™ PCR buffer (Boehringer Mannheim, Germany)
15   supplemented with 200 µM of each dNTP, 2.6 units of
HiFidelity™ Expand enzyme mix and 300 pmol of each primer:

#27184
5'-GCT GCA <u>GGA TCC</u> GTT TCA ATT TAT GTT CAA AGA TCT CCT GGA
20   GAG TAT CCA GCA TGG GAC CCA A-3'

#28495
5'-GC ACA AGC TT<u>G CGG CCG C</u>TA ATT GAG TGG TTC CCA CGG ACC G -
25   3'

(BamHI and NotI restriction sites underlined).

The primers were designed to amplify the CBD-encoding DNA of
30   the cellulase-encoding gene of *Bacillus agaradherens* NCIMB
40482.

The PCR reaction was performed using a DNA thermal cycler
(Landgraf, Germany). One incubation at 94°C for 2 min, 60°C

32

for 30 sec and 72°C for 45 sec was followed by ten cycles of
PCR performed using a cycle profile of denaturation at 94°C
for 30 sec, annealing at 60°C for 30 sec, and extension at
72°C for 45 sec and twenty cycles of denaturation at 94°C for
5 30 sec, 60°C for 30 sec and 72°C for 45 sec (at this
elongation step, 20 sec are added every cycle). 10 µl
aliquots of amplification product were analyzed by
electrophoresis in 1.5 % agarose gels (NuSieve™, FMC) with
ReadyLoad™ 100bp DNA ladder (GibcoBRL, Denmark) as a size
10 marker.

Cloning by polymerase chain reaction (PCR):
Subcloning of PCR fragments
40 µl aliquots of PCR products generated as described above
15 were purified using QIAquick™ PCR purification kit (Qiagen,
USA) according to the manufacturer's instructions. The
purified DNA was eluted in 50 µl of 10mM Tris-HCl, pH 8.5. 25
µl of the purified PCR fragment was digested with BamHI and
NotI, subjected to electrophoresis in 1.5% low gelling
20 temperature agarose (SeaPlaque™ GTG, FMC) gels, and the
relevant fragment was excised from the gels and purified
using QIAquick™ Gel extraction kit (Qiagen, USA) according
to the manufacturer's instructions. The isolated DNA fragment
was then ligated to BamHI-NotI digested pMB335, and the
25 ligation mixture was used to transform *E. coli* SJ2.

Identification and characterization of positive clones
Cells were plated on LB agar plates containing Ampicillin
(200 µg/ml) and incubated at 37°C overnight. The next day,
30 colonies were restreaked onto fresh LB-Ampicillin agar plates
and incubated at 37°C overnight. The following day, single
colonies were transferred to liquid LB medium containing

33

Ampicillin (200 μg/ml) and incubated overnight at 37°C with
shaking at 250 rpm.

Plasmids were extracted from the liquid cultures using QIAgen
5   Plasmid Purification mini kit (Qiagen, USA) according to the
manufacturer's instructions. 5 μl samples of the plasmids
were digested with BamHI and NotI. The digestions were
checked by gel electrophoresis on a 1.5% agarose gel (NuSie-
ve™, FMC). The appearance of a DNA fragment of the same
10   size as seen from the PCR amplification indicated a positive
clone.

One positive clone, containing the fusion construct of the
Termamyl™ α-amylase gene and the CBD of *Bacillus*
15   *agaradherens* NCIMB 40482 alkaline cellulase Cel5A, was
designated MBamyC5A.

**Cloning of the fusion construct into a *Bacillus*-based
expression vector**
20   As mentioned previously, the amyL gene of *B. licheniformis*
(contained in the pDN1528 vector) is actively expressed in *B.
subtilis*, resulting in the production of active α-amylase
appearing in the supernatant. For expression purposes, the
DNA encoding the fusion protein as constructed above was
25   introduced to pDN1528. This was done by digesting pMBamyC5A
and pDN1528 with SalI-NotI, purifying the fragments and
ligating the 4.7 kb pDN1528 SalI-NotI fragment with the 0.5
kb pMBamyC5A SalI-NotI fragment. This created an inframe
fusion of the hybrid construction with the Termamyl™ gene.
30   The DNA sequence of the fusion construction of pMB378, and
the corresponding amino acid sequence, are shown in SEQ ID
No. 5 and SEQ ID No. 6, respectively.

34

The ligation mixture was used to transform competent cells of
B. subtilis PL2306. Cells were plated on LB agar plates
containing chloramphenicol (6 μg/ml), 0.4% glucose and 10mM
potassium hydrogen phosphate, and incubated at 37°C
5   overnight. The next day, colonies were restreaked onto fresh
LBPG chloramphenicol agar plates and incubated at 37°C
overnight. The following day, single colonies of each clone
were transferred to liquid LB medium containing
chloramphenicol (6 μg/ml) and incubated overnight at 37°C
10   with shaking at 250 rpm.

Plasmids were extracted from the liquid cultures using QIAgen
Plasmid Purification mini kit (Qiagen, USA) according to the
manufacturer's instructions. However, the resuspension buffer
15   was supplemented with 1 mg/ml of chicken egg white lysozyme
(SIGMA, USA) prior to lysing the cells at 37°C for 15
minutes. 5 μl samples of the plasmids were digested with
BamHI and NotI. The digestions were checked by gel
electrophoresis on a 1.5% agarose gel (NuSieve™, FMC). The
20   appearance of a DNA fragment of the same size as seen from
the PCR amplification indicated a positive clone. One
positive clone was designated MB378.

Expression, secretion and functional analysis of the fusion
25   **protein**

The clone MB378 (expressing Termamyl™ fused to *Bacillus
agaradherens* Cel5A CBD) was incubated for 20 hours in SB
medium at 37°C with shaking at 250 rpm. 1 ml of cell-free
supernatant was mixed with 200 μl of 10% Avicel™. The
30   mixture was incubated for 1 hour at 0°C and then centrifuged
for 5 minutes at 5000 x g. The pellet was resuspended in 100
μl of SDS-PAGE buffer, and the suspension was boiled at 95°C
for 5 minutes, centrifuged at 5000 x g for 5 minutes, and 25

μl was loaded onto a 4-20% Laemmli Tris-Glycine, SDS-PAGE
NOVEX™ gel (Novex, USA). The samples were subjected to
electrophoresis in an Xcell™ Mini-Cell (NOVEX, USA) as
recommended by the manufacturer. All subsequent handling of
5   gels, including staining (Coomassie), destaining and drying,
were performed as described by the manufacturer.

The appearance of a protein band of molecular weight approx.
60 kDa indicated expression in *B. subtilis* of the Termamyl™-
10 CBD fusion encoded on the plasmid pMB378.

EXAMPLE 3

15 This example describes fusion of Termamyl™ and the CBD from
Cellulomonas fimi (ATCC484) *cenA* gene using the sequence
overlap extension (SOE) procedure [see, e.g., Sambrook et
al., Ausubel et al., or C.R. Harwood and S.M. Cutting (*loc.
cit.*)]. The final construction is as follows: Termamyl™
20 promoter - Termamyl™ signal peptide - cenA CBD - linker -
mature Termamyl™.

Amplification of the Termamyl™ fragment for SOE
Approximately 10-20 ng of plasmid pDN1528 was PCR amplified
25 in HiFidelity™ PCR buffer (Boehringer Mannheim, Germany)
supplemented with 200 μM of each dNTP, 2.6 units of
HiFidelity™ Expand enzyme mix, and 100 pmol of each primer:

#4576
30 5'-CTC GTC CCA ATC GGT TCC GTC -3´

#28403

5'-TGC ACT GGT ACA GTT CCT ACA ACT AGT CCT ACA CGT <u>GCA AAT
CTT AAT GGG ACG CTG</u> -3'

The part of the primer #28403 constituting a fragment of the
5   Termamyl™ sequence is underlined. The sequence on the 5´
side of this underlined sequence is that coding for the
linker region to the CBD.

The PCR reaction was performed using a DNA thermal cycler
10   (Landgraf, Germany). One incubation at 94°C for 2 min, 55°C
for 30 sec and 72°C for 45 sec was followed by twenty cycles
of PCR performed using a cycle profile of denaturation at
96°C for 10 sec, annealing at 55°C for 30 sec, and extension
at 72°C for 45 sec. 10 µl aliquots of the amplification pro-
15   duct were analyzed by electrophoresis in 1.0 % agarose gels
(NuSieve™, FMC) with ReadyLoad™ 100bp DNA ladder (GibcoBRL,
Denmark) as a size marker.

40 µl aliquots of the PCR product generated as described
20   above were purified using QIAquick™ PCR purification kit
(Qiagen, USA) according to the manufacturer's instructions.
The purified DNA was eluted in 50 µl of 10mM Tris-HCl, pH
8.5.

25   **Isolation of genomic DNA**
Cellulomonas fimi ATCC484 was grown in TY medium at 30°C with
shaking at 250 rpm for 24 hours. Cells were harvested by
centrifugation.

30   Genomic DNA was isolated as described by Pitcher et al.,
<u>Lett. Appl. Microbiol.</u> <u>8</u> (1989), pp. 151-156.

In vitro **amplification of the CBD of** *Cellulomonas fimi*

37

(ATCC484) *cenA* gene for SOE procedure
Approximately 100-200 ng of genomic DNA was PCR amplified in
HiFidelity™ PCR buffer (Boehringer Mannheim, Germany)
supplemented with 200 µM of each dNTP, 2.6 units of
5   HiFidelity™ Expand enzyme mix, and 100 pmol of each primer:

#8828
5'-CTG CCT CAT TCT GCA GCA GCG GCG GCA AAT CTT AAT GCT CCC
GGC TGC CGC GTC GAC TAC -3'
10

#28404
5'-TGT AGG AAC TGT ACC AGT GCA CGT GGT GCC GTT GAG C -3'

(PstI restriction site underlined).
15

The primers were designed to amplify the DNA encoding the
cellulose-binding domain of the CenA-encoding gene of
Cellulomonas fimi (ATCC484). The DNA sequence was extracted
from the database GenBank under the accession number M15823.
20

PCR cycling was performed as follows: One incubation at 94°C
for 2 min, 55°C for 30 sec and 72°C for 45 sec was followed
by thirty cycles of PCR performed using a cycle profile of
denaturation at 96°C for 10 sec, annealing at 55°C for 30
25   sec, and extension at 72°C for 45 sec. 10 µl aliquots of the
amplification product were analyzed by electrophoresis in 1.0
% agarose gels (NuSieve™, FMC) with ReadyLoad™ 100bp DNA
ladder (GibcoBRL, Denmark) as a size marker.

30   40 µl aliquots of the PCR product generated as described
above were purified using QIAquick™ PCR purification kit
(Qiagen, USA) according to the manufacturer's instructions.
The purified DNA was eluted in 50 µl of 10mM Tris-HCl, pH

38

8.5.

SOE of the CBD from *Cellulomonas fimi* (ATCC484) cenA gene and the Termamyl™ gene

5   Approximately 100-200 ng of the PCR amplified Termamyl™
     fragment and the PCR amplified *cenA* CBD fragment were used in
     a second round of PCR. SOE of the two fragments was performed
     in
     in HiFidelity™ PCR buffer (Boehringer Mannheim, Germany)
10   supplemented with 200 µM of each dNTP, 2.6 units of
     HiFidelity™ Expand enzyme mix.

     A touch-down PCR cycling was performed as follows: One
     incubation at 96°C for 2 min, 60°C for 2 min and 72°C for 45
15   sec. This cycle was repeated ten times with a 1°C decrease of
     the annealing temperature at each cycle.

     A third PCR reaction was started by adding 100 pmol of the
     two flanking primers #8828 and #4576 (*vide supra*) to amplify
20   the hybrid DNA. PCR was performed by incubating the SOE
     reaction mixture at 96°C for 2 min, 55°C for 30 sec and 72°C
     for 45 sec. This was followed by twenty cycles of PCR
     performed using a cycle profile of denaturation at 96°C for
     10 sec, annealing at 55°C for 30 sec, and extension at 72°C
25   for 45 sec. 10 µl aliquots of the amplification product were
     analyzed by electrophoresis in 1.0 % agarose gels (NuSieve™,
     FMC) with ReadyLoad™ 100bp DNA ladder (GibcoBRL, Denmark) as
     a size marker. The SOE fragment had the expected size of 879
     bp.

30

Subcloning of the SOE fragment coding for the CBD-Termamyl hybrid

39

40 µl of the SOE-PCR product generated as described above was
purified using QIAquick™ PCR purification kit (Qiagen, USA)
according to the manufacturer's instructions. The purified D-
NA was eluted in 50 µl of 10mM Tris-HCl, pH 8.5. 25 µl of the
5   purified PCR fragment was digested with PstI and KpnI,
subjected to electrophoresis in 1.0% low gelling temperature
agarose (SeaPlaque™ GTG, FMC) gels, and a fragment of 837 bp
was excised from the gel and purified using QIAquick™ Gel
extraction Kit (Qiagen, USA) according to the manufacturer's
10   instructions. The isolated DNA fragment was then ligated to
PstI- and KpnI-digested pDN1981, and the ligation mixture was
used to transform competent cells of *B. subtilis* PL2306.
Cells were plated on LB agar plates containing Kanamycin (10
µg/ml), 0.4% glucose and 10mM potassium hydrogen phosphate,
15   and incubated at 37°C overnight. The next day, colonies were
restreaked onto fresh LBPG Kanamycin agar plates and in-
cubated at 37°C overnight. The following day, single colonies
of each clone were transferred to liquid LB medium containing
Kanamycin (10 µg/ml) and incubated overnight at 37°C with
20   shaking at 250 rpm.

Plasmids were extracted from the liquid cultures using QIAgen
Plasmid Purification mini kit (Qiagen, USA) according to the
manufacturer's instructions. However, the resuspension buffer
25   was supplemented with 1 mg/ml of chicken egg white lysozyme
(SIGMA, USA) prior to lysing the cells at 37°C for 15
minutes. 5 µl samples of the plasmids were digested with PstI
and KpnI. The digestions were checked by gel electrophoresis
on a 1.5% agarose gel (NuSieve™, FMC). The appearance of a
30   DNA fragment of 837 bp, the same size as seen from the PCR
amplification, indicated a positive clone. One positive clone
was designated MOL1297.

**Expression, secretion and functional analysis of the fusion
protein**

The clone MOL1297 (expressing *C. fimi cenA* CBD fused to the

5   N-terminal of Termamyl™) was incubated for 20 hours in SB
medium at 37°C with shaking at 250 rpm. 1 ml of cell-free
supernatant was mixed with 200 µl of 10% Avicel™. The
mixture was incubated for 1 hour at 0°C and then centrifuged
for 5 min at 5000 x g. The pellet was resuspended in 100 µl

10   of SDS-PAGE buffer, boiled at 95°C for 5 minutes, centrifuged
at 5000 x g for 5 minutes, and 25 µl was loaded on a 4-20%
Laemmli Tris-Glycine, SDS-PAGE NOVEX gel (Novex, USA). The
samples were subjected to electrophoresis in an Xcell™ Mini-
Cell (NOVEX, USA) as recommended by the manufacturer. All

15   subsequent handling of gels including staining (Coomassie),
destaining and drying, was performed as described by the
manufacturer.

The appearance of a protein band of MW approx. 85 kDa

20   indicated expression in *B. subtilis* of the CBD-Termamyl™
fusion.

The encoding sequence for the *C. fimi cenA* CBD-Termamyl
hybrid is shown in SEQ ID No. 7 (in which lower case letters

25   indicate the CBD-encoding part of the sequence). The
corresponding amino acid sequence of the hybrid is shown in
SEQ ID No. 8 (in which lower case letters indicate the CBD
amino acid sequence).

30   **EXAMPLE 4**

This example describes the construction of fusion proteins
(enzyme hybrid) from a lipase (Lipolase™; *Humicola*

41

lanuginosa lipase) and a CBD. A construction with an N-terminal CBD was chosen, since the N-terminal of the enzyme is far from the active site, whereas the C-terminal is in relatively close proximity to the active site.

5

pIVI450 construction (CBD-linker-lipase)
This construct was made in order to express a protein having the *Myceliophthora thermophila* cellulase CBD and linker at the N-terminal of Lipolase™.

10

A PCR fragment was created using the clone pA2C161 (DSM 9967) containing the *M. thermophila* cellulase gene as template, and the following oligomers as primers:

15   #8202
      5' ACGTAGTGGCCACGCTAGGCGAGGTGGTGG 3'

      #19672
      5' CCACACTTCTCTTCCTTCCTC 3'

20

The PCR fragment was cut with BamHI and BalI, and cloned into pAHL which was also cut with BamHI and BalI just upstream of the presumed signal peptide processing site. The cloning was verified by sequencing (see SEQ ID No. 9).

25

Removing linker between CBD and lipase
This construct is made so that any linker of interest can be inserted between the CBD and the lipase in order to find an optimal linker.

30

An NheI site is introduced by the USE technique (Stratagene catalogue No. 200509) between the CBD and linker region in pIVI450, creating pIVI450+NheI site. pIVI450+NheI site is cut with XhoI and NheI, isolating the vector containing the

42

CBD part.

The plasmid pIVI392 is cut with XhoI and NheI, and the
fragment containing the Lipolase™ gene (minus signal peptide
5   encoding sequence) is isolated.

The DNA fragments are ligated, generating pIVI450 CBD-NheI
site-Lipolase™ containing an NheI site between the CBD and
the lipase gene. In this NheI site different linkers can be
10   introduced.

Introduction of non-glycosylated linker
The protein expressed from the construct described here
contains a construction of the type:
15   CBD-nonglycosylated linker-lipase.

The amino acid sequence of the linker is as follows:

NNNPQQGNPNQGGNNGGGNQGGGNGG
20

PCR is performed with the following primers:

#29315
5' GATCTAGCTAGCAACAATAACCCCCAGCAGGGCAACCCCAACCAGGGC
25   GGGAACAACGGC 3'

#29316
5' GATCTAGCTAGCGCCGCCGTTGCCGCCGCCCTGGTTGCCGCCGCCGTT
GTTCCCGCCCTG 3'
30

The PCR fragment is cut with NheI, the vector pIVI450 CBD-
NheI-Lipolase™ is likewise cut with NheI, and the two
fragments are ligated, creating:
pIVI450 CBD-Nonglycosylated linker-Lipolase™ (SEQ ID No.

43

10).

Introduction of *H. insolens* family 45 cellulase linker
The protein expressed from the construct described here
5   contains a construction of the type:
CBD-glycosylated linker-lipase.

The amino acid sequence of the linker is as follows:

10 VQIPSSSTSSPVNQPTSTSTTSTSTTSSPPVQPTTPS

PCR is performed with the following primers:

#29313
15 5' GATACTGCTAGCGTCCAGATCCCCTCCAGC 3'

#29314
5' GATACTGCTAGCGCTGGGAGTCGTAGGCTG 3'

20 The PCR fragment is cut with NheI, the vector pIVI450 CBD-
NheI-Lipolase™ is likewise cut with NheI, and the two
fragments are ligated, creating:
pIVI450 CBD-*H. insolens* family 45 cellulase linker-Lipolase™
(SEQ ID No. 11).

44

SEQUENCE LISTING

(1) GENERAL INFORMATION:
5

        (i) APPLICANT:
            (A) NAME: NOVO NORDISK A/S
            (B) STREET: Novo Alle
            (C) CITY: Bagsvaerd
10   (E) COUNTRY: Denmark
            (F) POSTAL CODE (ZIP): DK-2880
            (G) TELEPHONE: +45 44 44 88 88
            (H) TELEFAX: +45 44 49 32 56

15   (ii) TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING
                OR STAINS FROM CELLULOSIC FABRIC

        (iii) NUMBER OF SEQUENCES: 6

20   (iv) COMPUTER READABLE FORM:
            (A) MEDIUM TYPE: Floppy disk
            (B) COMPUTER: IBM PC compatible
            (C) OPERATING SYSTEM: PC-DOS/MS-DOS
            (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
25

    (2) INFORMATION FOR SEQ ID NO: 1:

        (i) SEQUENCE CHARACTERISTICS:
30   (A) LENGTH: 2253 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

35   (ii) MOLECULE TYPE: DNA (genomic)

40   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ATGAAACAAC AAAAACGGCT TTACGCCCGA TTGCTGACGC TGTTATTTGC GCTCATCTTC      60

TTGCTGCCTC ATTCTGCAGC AGCGGCGGCA AATCTTAATG GGACGCTGAT GCAGTATTTT     120

GAATGGTACA TGCCCAATGA CGGCCAACAT TGGAAGCGTT TGCAAAACGA CTCGGCATAT     180

TTGGCTGAAC ACGGTATTAC TGCCGTCTGG ATTCCCCCGG CATATAAGGG AACGAGCCAA     240

GCGGATGTGG GCTACGGTGC TTACGACCTT TATGATTTAG GGGAGTTTCA TCAAAAGGG     300

ACGGTTCGGA CAAAGTACGG CACAAAAGGA GAGCTGCAAT CTGCGATCAA AAGTCTTCAT     360

TCCCGCGACA TTAACGTTTA CGGGGATGTG GTCATCAACC ACAAAGGCGG CGCTGATGCG     420

ACCGAAGATG TAACCGCGGT TGAAGTCGAT CCCGCTGACC GCAACCGCGT AATCTCAGGA     480

GAACACCTAA TTAAAGCCTG GACACATTTT CATTTTCCGG GGGCCGGCAG CACATACAGC     540

GATTTTAAAT GGCATTGGTA CCATTTTGAC GGAACCGATT GGGACGAGTC CCGAAAGCTG     600

AACCGCATCT ATAAGTTTCA AGGAAAGGCT TGGGATTGGG AAGTTTCCAA TGAAAACGGC     660

AACTATGATT ATTTGATGTA TGCCGACATC GATTATGACC ATCCTGATGT CGCAGCAGAA     720

ATTAAGAGAT GGGGCACTTG GTATGCCAAT GAACTGCAAT TGGACGGAAA CCGTCTTGAT     780

GCTGTCAAAC ACATTAAATT TTCTTTTTTG CGGGATTGGG TTAATCATGT CAGGGAAAAA     840

ACGGGGAAGG AAATGTTTAC GGTAGCTGAA TATTGGCAGA ATGACTTGGG CGCGCTGGAA     900

AACTATTTGA ACAAAACAAA TTTTAATCAT TCAGTGTTTG ACGTGCCGCT TCATTATCAG     960

TTCCATGCTG CATCGACACA GGGAGGCGGC TATGATATGA GGAAATTGCT GAACGGTACG    1020

GTCGTTTCCA AGCATCCGTT GAAATCGGTT ACATTTGTCG ATAACCATGA TACACAGCCG    1080

GGGCAATCGC TTGAGTCGAC TGTCCAAACA TGGTTTAAGC CGCTTGCTTA CGCTTTTATT    1140

CTCACAAGGG AATCTGGATA CCCTCAGGTT TTCTACGGGG ATATGTACGG GACGAAAGGA    1200

GACTCCCAGC GCGAAATTCC TGCCTTGAAA CACAAAATTG AACCGATCTT AAAAGCGAGA    1260
```

46

```
AAACAGTATG CGTACGGAGC ACAGCATGAT TATTTCGACC ACCATGACAT TGTCGGCTGG      1320

ACAAGGGAAG GCGACAGCTC GGTTGCAAAT TCAGGTTTGG CGGCATTAAT AACAGACGGA      1380

CCCGGTGGGG CAAAGCGAAT GTATGTCGGC CGGCAAAACG CCGGTGAGAC ATGGCATGAC      1440

ATTACCGGAA ACCGTTCGGA GCCGGTTGTC ATCAATTCGG AAGGCTGGGG AGAGTTTCAC      1500

GTAAACGGCG GATCCGTTTC AATTTATGTT CAAAGATCTG GCGGACCTGG AACGCCAAAT      1560

AATGGCAGAG GAATTGGTTA TATTGAAAAT GGTAATACCG TAACTTACAG CAATATAGAT      1620

TTTGGTAGTG GTGCAACAGG GTTCTCTGCA ACTGTTGCAA CGGAGGTTAA TACCTCAATT      1680

CAAATCCGTT CTGACAGTCC TACCGGAACT CTACTTGGTA CCTTATATGT AAGTTCTACC      1740

GGCAGCTGGA ATACATATCA ACCGTATCTA CAAACATCAG CAAAATTACC GGCGTTCATG      1800

ATATTGTATT GGTATTCTCA GGTCCAGTCA ATGTGGACAA CTTCATATTT AGCAGAAGTT      1860

CACCAGTGCC TGCACCTGGT GATAACACAA GAGACGCATA TTCTATCATT CAGGCCGAGG      1920

ATTATGACAG CAGTTATGGT CCCAACCTTC AAATCTTTAG CTTACCAGGT GGTGGCAGCG      1980

CTTGGCTATA TTGAAAATGG TTATTCCACT ACCTATAAAA ATATTGATTT TGGTGACGGC      2040

GCAACGTCCG TAACAGCAAG AGTAGCTACC CAGAATGCTA CTACCATTCA GGTAAGATTG      2100

GGAAGTCCAT CGGGTACATT ACTTGGAACA ATTTACGTGG GGTCCACAGG AAGCTTTGAT      2160

ACTTATAGGG ATGTATCCGC TACCATTAGT AATACTGCGG GTGTAAAAGA TATTGTTCTT      2220

GTATTCTCAG GTCCTGTTAA TGTTGACTGG TAG                                   2253
```

(2) INFORMATION FOR SEQ ID NO: 2:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 750 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: single

47

      (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: protein

5

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

10
```
Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe
1               5                   10                  15

Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ala Ala Asn Leu
            20                  25                  30

Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro Asn Asp Gly
        35                  40                  45

Gln His Trp Lys Arg Leu Gln Asn Asp Ser Ala Tyr Leu Ala Glu His
        50                  55                  60

Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln
65                  70                  75                  80

Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe
                85                  90                  95

His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu Leu
            100                 105                 110

Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr Gly
            115                 120                 125

Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp Val
        130                 135                 140

Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser Gly
145                 150                 155                 160

Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Ala Gly
                165                 170                 175
```

48

```
        Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly Thr
                    180                 185                 190

  5     Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln Gly
                    195                 200                 205

        Lys Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn Tyr Asp Tyr
                210                 215                 220
 10
        Leu Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val Ala Ala Glu
            225                 230                 235                 240

        Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp Gly
 15                     245                 250                 255

        Asn Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp
                    260                 265                 270

 20     Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr Val
                    275                 280                 285

        Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu Asn
                290                 295                 300
 25
        Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr Gln
            305                 310                 315                 320

        Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys Leu
 30                         325                 330                 335

        Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser Val Thr Phe
                    340                 345                 350

 35     Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr Val
                    355                 360                 365

        Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Glu
                370                 375                 380
 40
        Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys Gly
```

49

```
          385                    390                    395                    400

        Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro Ile
                        405                    410                    415
   5
        Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr Phe
                        420                    425                    430

        Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser Val
   10                   435                    440                    445

        Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ala
                    450                    455                    460

   15   Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His Asp
        465                    470                    475                    480

        Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly Trp
                            485                    490                    495
   20
        Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Arg
                        500                    505                    510

        Ser Gly Gly Pro Gly Thr Pro Asn Asn Gly Arg Gly Ile Gly Tyr Ile
   25           515                    520                    525

        Glu Asn Gly Asn Thr Val Thr Tyr Ser Asn Ile Asp Phe Gly Ser Gly
                530                    535                    540

   30   Ala Thr Gly Phe Ser Ala Thr Val Ala Thr Glu Val Asn Thr Ser Ile
        545                    550                    555                    560

        Gln Ile Arg Ser Asp Ser Pro Thr Gly Thr Leu Leu Gly Thr Leu Tyr
                        565                    570                    575
   35
        Val Ser Ser Thr Gly Ser Trp Asn Thr Tyr Gln Pro Tyr Leu Gln Thr
                    580                    585                    590

        Ser Ala Lys Leu Pro Ala Phe Met Ile Leu Tyr Trp Tyr Ser Gln Val
   40           595                    600                    605
```

50

```
Gln Ser Met Trp Thr Thr Ser Tyr Leu Ala Glu Val His Gln Cys Leu
    610                 615                 620

His Leu Val Ile.Thr Gln Glu Thr His Ile Leu Ser Phe Arg Pro Arg
625                 630                 635                 640

Ile Met Thr Ala Val Met Val Pro Thr Phe Lys Ser Leu Ala Tyr Gln
            645                 650                 655

Val Val Ala Ala Leu Gly Tyr Ile Glu Asn Gly Tyr Ser Thr Thr Tyr
            660                 665                 670

Lys Asn Ile Asp Phe Gly Asp Gly Ala Thr Ser Val Thr Ala Arg Val
        675                 680                 685

Ala Thr Gln Asn Ala Thr Thr Ile Gln Val Arg Leu Gly Ser Pro Ser
    690                 695                 700

Gly Thr Leu Leu Gly Thr Ile Tyr Val Gly Ser Thr Gly Ser Phe Asp
705                 710                 715                 720

Thr Tyr Arg Asp Val Ser Ala Thr Ile Ser Asn Thr Ala Gly Val Lys
            725                 730                 735

Asp Ile Val Leu Val Phe Ser Gly Pro Val Asn Val Asp Trp
            740                 745                 750
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1203 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

51

```
    ATGAAAAAGA TAACTACTAT TTTTGTCGTA TTGCTTATGA CAGTGGCGTT GTTCAGTATA    60

    GGAAACACGA CTGCTGCTGA TAATGATTCA GTTGTAGAAG AACATGGGCA ATTAAGTATT   120
 5
    AGTAACGGTG AATTAGTCAA TGAACGAGGC GAACAAGTTC AGTTAAAAGG GATGAGTTCC   180

    CATGGTTTGC AATGGTACGG TCAATTTGTA AACTATGAAA GTATGAAATG GCTAAGAGAT   240

10  GATTGGGGAA TAAATGTATT CCGAGCAGCA ATGTATACCT CTTCAGGAGG ATATATTGAT   300

    GATCCATCAG TAAAGGAAAA AGTAAAAGAG GCTGTTGAAG CTGCGATAGA CCTTGATATA   360

    TATGTGATCA TTGATTGGCA TATCCTTTCA GACAATGACC CAAATATATA TAAAGAAGAA   420
15
    GCGAAGGATT TCTTTGATGA AATGTCAGAG TTGTATGGAG ACTATCCGAA TGTGATATAC   480

    GAAATTGCAA ATGAACCGAA TGGTAGTGAT GTTACGTGGG CAATCAAAT AAAACCGTAT    540

20  GCAGAGGAAG TCATTCCGAT TATTCGTAAC AATGACCCTA ATAACATTAT TATTGTAGGT   600

    ACAGGTACAT GGAGTCAGGA TGTCCATCAT GCAGCTGATA ATCAGCTTGC AGATCCTAAC   660

    GTCATGTATG CATTTCATTT TTATGCAGGG ACACATGGTC AAAATTTACG AGACCAAGTA   720
25
    GATTATGCAT TAGATCAAGG AGCAGCGATA TTTGTTAGTG AATGGGGAAC AAGTGCAGCT   780

    ACAGGTGATG GTGGCGTGTT TTTAGATGAA GCACAAGTGT GGATTGACTT TATGGATGAA   840

30  AGAAATTTAA GCTGGGCCAA CTGGTCTCTA ACGCATAAAG ATGAGTCATC TGCAGCGTTA   900

    ATGCCAGGTG CAAATCCAAC TGGTGGTTGG ACAGAGGCTG AACTATCTCC ATCTGGTACA   960

    TTTGTGAGGG AAAAAATAAG AGAATCAGCA TCTATTCCGC CAAGCGATCC AACACCGCCA  1020
35
    TCTGATCCAG GAGAACCGGA TCCAACGCCC CCAAGTGATC CAGGAGAGTA TCCAGCATGG  1080

    GATCCAAATC AAATTTACAC AAATGAAATT GTGTACCATA ACGGCCAGCT ATGGCAAGCA  1140

40  AAATGGTGGA CACAAAATCA AGAGCCAGGT GACCCGTACG GTCCGTGGGA ACCACTCAAT  1200
```

52

TAA   1203

(2) INFORMATION FOR SEQ ID NO: 4:

5   (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 400 amino acids
               (B) TYPE: amino acid
               (C) STRANDEDNESS: single
               (D) TOPOLOGY: linear
10

           (ii) MOLECULE TYPE: protein

15

           (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

        Met Lys Lys Ile Thr Thr Ile Phe Val Val Leu Leu Met Thr Val Ala
        1   5   10   15
20

        Leu Phe Ser Ile Gly Asn Thr Thr Ala Ala Asp Asn Asp Ser Val Val
                    20   25   30

        Glu Glu His Gly Gln Leu Ser Ile Ser Asn Gly Glu Leu Val Asn Glu
25   35   40   45

        Arg Gly Glu Gln Val Gln Leu Lys Gly Met Ser Ser His Gly Leu Gln
                50   55   60

30   Trp Tyr Gly Gln Phe Val Asn Tyr Glu Ser Met Lys Trp Leu Arg Asp
                65   70   75   80

        Asp Trp Gly Ile Asn Val Phe Arg Ala Ala Met Tyr Thr Ser Ser Gly
                    85   90   95
35

        Gly Tyr Ile Asp Asp Pro Ser Val Lys Glu Lys Val Lys Glu Ala Val
                    100   105   110

        Glu Ala Ala Ile Asp Leu Asp Ile Tyr Val Ile Ile Asp Trp His Ile
40   115   120   125

53

```
      Leu Ser Asp Asn Asp Pro Asn Ile Tyr Lys Glu Glu Ala Lys Asp Phe
          130                 135             140

  5   Phe Asp Glu Met Ser Glu Leu Tyr Gly Asp Tyr Pro Asn Val Ile Tyr
      145                 150             155             160

      Glu Ile Ala Asn Glu Pro Asn Gly Ser Asp Val Thr Trp Gly Asn Gln
                      165             170             175

 10   Ile Lys Pro Tyr Ala Glu Glu Val Ile Pro Ile Ile Arg Asn Asn Asp
                  180             185             190

      Pro Asn Asn Ile Ile Ile Val Gly Thr Gly Thr Trp Ser Gln Asp Val
              195             200             205

 15
      His His Ala Ala Asp Asn Gln Leu Ala Asp Pro Asn Val Met Tyr Ala
              210             215             220

 20   Phe His Phe Tyr Ala Gly Thr His Gly Gln Asn Leu Arg Asp Gln Val
      225             230             235             240

      Asp Tyr Ala Leu Asp Gln Gly Ala Ala Ile Phe Val Ser Glu Trp Gly
                  245             250             255

 25   Thr Ser Ala Ala Thr Gly Asp Gly Gly Val Phe Leu Asp Glu Ala Gln
                  260             265             270

      Val Trp Ile Asp Phe Met Asp Glu Arg Asn Leu Ser Trp Ala Asn Trp
              275             280             285

 30
      Ser Leu Thr His Lys Asp Glu Ser Ser Ala Ala Leu Met Pro Gly Ala
          290             295             300

 35   Asn Pro Thr Gly Gly Trp Thr Glu Ala Glu Leu Ser Pro Ser Gly Thr
      305             310             315             320

      Phe Val Arg Glu Lys Ile Arg Glu Ser Ala Ser Ile Pro Pro Ser Asp
                  325             330             335

 40   Pro Thr Pro Pro Ser Asp Pro Gly Glu Pro Asp Pro Thr Pro Pro Ser
                  340             345             350
```

54

        Asp Pro Gly Glu Tyr Pro Ala Trp Asp Pro Asn Gln Ile Tyr Thr Asn
                355   360   365

5   Glu Ile Val Tyr His Asn Gly Gln Leu Trp Gln Ala Lys Trp Trp Thr
                370   375   380

        Gln Asn Gln Glu Pro Gly Asp Pro Tyr Gly Pro Trp Glu Pro Leu Asn
        385   390   395   400
10

    (2) INFORMATION FOR SEQ ID NO: 5:

        (i) SEQUENCE CHARACTERISTICS:
15   (A) LENGTH: 1683 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

20   (ii) MOLECULE TYPE: DNA (genomic)

25   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

    ATGAAACAAC AAAAACGGCT TTACGCCCGA TTGCTGACGC TGTTATTTGC GCTCATCTTC   60

    TTGCTGCCTC ATTCTGCAGC AGCGGCGGCA AATCTTAATG GGACGCTGAT GCAGTATTTT   120
30
    GAATGGTACA TGCCCAATGA CGGCCAACAT TGGAAGCGTT TGCAAAACGA CTCGGCATAT   180

    TTGGCTGAAC ACGGTATTAC TGCCGTCTGG ATTCCCCCGG CATATAAGGG AACGAGCCAA   240

35 GCGGATGTGG GCTACGGTGC TTACGACCTT TATGATTTAG GGGAGTTTCA TCAAAAAGGG   300

    ACGGTTCGGA CAAAGTACGG CACAAAAGGA GAGCTGCAAT CTGCGATCAA AAGTCTTCAT   360

    TCCCGCGACA TTAACGTTTA CGGGGATGTG GTCATCAACC ACAAAGGCGG CGCTGATGCG   420
40
    ACCGAAGATG TAACCGCGGT TGAAGTCGAT CCCGCTGACC GCAACCGCGT AATCTCAGGA   480

55

```
     GAACACCTAA TTAAAGCCTG GACACATTTT CATTTTCCGG GGGCCGGCAG CACATACAGC      540

     GATTTTAAAT GGCATTGGTA CCATTTTGAC GGAACCGATT GGGACGAGTC CCGAAAGCTG      600
 5
     AACCGCATCT ATAAGTTTCA AGGAAAGGCT TGGGATTGGG AAGTTTCCAA TGAAAACGGC      660

     AACTATGATT ATTTGATGTA TGCCGACATC GATTATGACC ATCCTGATGT CGCAGCAGAA      720

10   ATTAAGAGAT GGGGCACTTG GTATGCCAAT GAACTGCAAT TGGACGGAAA CCGTCTTGAT      780

     GCTGTCAAAC ACATTAAATT TTCTTTTTTG CGGGATTGGG TTAATCATGT CAGGGAAAAA      840

     ACGGGGAAGG AAATGTTTAC GGTAGCTGAA TATTGGCAGA ATGACTTGGG CGCGCTGGAA      900
15
     AACTATTTGA ACAAAACAAA TTTTAATCAT TCAGTGTTTG ACGTGCCGCT TCATTATCAG      960

     TTCCATGCTG CATCGACACA GGGAGGCGGC TATGATATGA GGAAATTGCT GAACGGTACG     1020

20   GTCGTTTCCA AGCATCCGTT GAAATCGGTT ACATTTGTCG ATAACCATGA TACACAGCCG     1080

     GGGCAATCGC TTGAGTCGAC TGTCCAAACA TGGTTTAAGC CGCTTGCTTA CGCTTTTATT     1140

     CTCACAAGGG AATCTGGATA CCCTCAGGTT TTCTACGGGG ATATGTACGG GACCAAGGA      1200
25
     GACTCCCAGC GCGAAATTCC TGCCTTGAAA CACAAAATTG AACCGATCTT AAAGCGAGA      1260

     AAACAGTATG CGTACGGAGC ACAGCATGAT TATTTCGACC ACCATGACAT TGTCGGCTGG     1320

30   ACAAGGGAAG CGACAGCTC GGTTGCAAAT TCAGGTTTGG CGGCATTAAT AACAGACGGA      1380

     CCCGGTGGGG CAAAGCGAAT GTATGTCGGC CGGCAAAACG CCGGTGAGAC ATGGCATGAC     1440

     ATTACCGGAA ACCGTTCGGA GCCGGTTGTC ATCAATTCGG AAGGCTGGGG AGAGTTTCAC     1500
35
     GTAAACGGCG GATCCGTTTC AATTTATGTT CAAAGATCTC CTGGAGAGTA TCCAGCATGG     1560

     GATCCAAATC AAATTTACAC AAATGAAATT GTGTACCATA ACGGCCAGCT ATGGCAAGCA     1620

40   AAATGGTGGA CACAAAATCA AGAGCCAGGT GACCCGTACG GTCCGTGGGA ACCACTCAAT     1680
```

56

TAA   1683

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 560 amino acids
    (B) TYPE: amino acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe
1   5   10   15

Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ala Ala Ala Asn Leu
            20   25   30

Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro Asn Asp Gly
        35   40   45

Gln His Trp Lys Arg Leu Gln Asn Asp Ser Ala Tyr Leu Ala Glu His
        50   55   60

Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln
65   70   75   80

Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe
                85   90   95

His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu Leu
            100   105   110

Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr Gly
            115   120   125

```
        Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp Val
            130             135             140

   5    Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser Gly
        145             150             155             160

        Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Ala Gly
                        165             170             175

  10    Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly Thr
                        180             185             190

        Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln Gly
                195             200             205

  15
        Lys Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn Tyr Asp Tyr
            210             215             220

  20    Leu Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val Ala Ala Glu
        225             230             235             240

        Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp Gly
                        245             250             255

  25    Asn Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp
                        260             265             270

        Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr Val
                275             280             285

  30    Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu Asn
            290             295             300

  35    Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr Gln
        305             310             315             320

        Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys Leu
                        325             330             335

  40    Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser Val Thr Phe
                        340             345             350
```

58

```
        Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr Val
                355             360             365

  5     Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Glu
                370             375             380

        Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys Gly
        385             390             395             400
 10
        Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro Ile
                        405             410             415

        Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr Phe
 15                     420             425             430

        Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser Val
                435             440             445

 20     Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ala
                450             455             460

        Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His Asp
        465             470             475             480
 25
        Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly Trp
                        485             490             495

        Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Arg
 30                     500             505             510

        Ser Pro Gly Glu Tyr Pro Ala Trp Asp Pro Asn Gln Ile Tyr Thr Asn
                515             520             525

 35     Glu Ile Val Tyr His Asn Gly Gln Leu Trp Gln Ala Lys Trp Trp Thr
                530             535             540

        Gln Asn Gln Glu Pro Gly Asp Pro Tyr Gly Pro Trp Glu Pro Leu Asn
        545             550             555             560
 40
```

59

SEQ ID No. 7:

```
      ATGAAACAACAAAAACGGCTTTACGCCCGATTGCTGACGCTGTTATTTGCGCTCATCTTCT
      TGCTGCCTCATTCTGCAGC
  5   AGCGGCGGCAAATCTTAATgctcccggctgccgcgtcgactacgccgtcaccaaccagtgg
      cccggcggcttcggcgcca
      acgtcacgatcaccaacctcggcgacccgtctcgtcgtggaagctcgactggacctacac
      cgcaggccagcggatccag
      cagctgtggaacggcaccgcgtcgaccaacggcggccaggtctccgtcaccagcctgccct
 10   ggaacggcagcatcccgac
      cggcggcacggcgtcgttcgggttcaacggctcgtgggccgggtccaacccgacgccggcg
      tcgttctcgctcaacggca
      ccacgtgcactggtacagttcctacaactagtcctacacgtGCAAATCTTAATGGGACGCT
      GATGCAGTATTTTGAATGG
 15   TACATGCCCAATGACGGCCAACATTGGAGGCGTTTGCAAAACGACTCGGCATATTTGGCTG
      AACACGGTATTACTGCCGT
      CTGGATTCCCCCGGCATATAAGGGAACGAGCCAAGCGGATGTGGGCTACGGTGCTTACGAC
      CTTTATGATTTAGGGGAGT
      TTCATCAAAAAGGGACGGTTCGGACAAAGTACGGCACAAAAGGAGAGCTGCAATCTGCGAT
 20   CAAAAGTCTTCATTCCCGC
      GACATTAACGTTTACGGGGATGTGGTCATCAACCACAAAGGCGGCGCTGATGCGACCGAAG
      ATGTAACCGCGGTTGAAGT
      CGATCCCGCTGACCGCAACCGCGTAATTTCAGGAGAACACCTAATTAAAGCCTGGACACAT
      TTTCATTTTCCGGGGCGCG
 25   GCAGCACATACAGCGATTTTAAATGGCATTGGTACCATTTTGACGGAACCGATTGGGACGA
      GTCCCGAAAGCTGAACCGC
      ATCTATAAGTTTCAAGGAAAGGCTTGGGATTGGGAAGTTTCCAATGAAAACGGCAACTATG
      ATTATTTGATGTATGCCGA
      CATCGATTATGACCATCCTGATGTCGCAGCAGAAATTAAGAGATGGGGCACTTGGTATGCC
 30   AATGAACTGCAATTGGACG
      GTTTCCGTCTTGATGCTGTCAAACACATTAAATTTTCTTTTTTGCGGGATTGGGTTAATCA
      TGTCAGGGAAAAAACGGGG
      AAGGAAATGTTTACGGTAGCTGAATATTGGCAGAATGACTTGGGCGCGCTGGAAAACTATT
      TGAACAAAACAAATTTTAA
 35   TCATTCAGTGTTTGACGTGCCGCTTCATTATCAGTTCCATGCTGCATCGACACAGGGAGGC
      GGCTATGATATGAGGAAAT
      TGCTGAACGGTACGGTCGTTTCCAAGCATCCGTTGAAATCGGTTACATTTGTCGATAACCA
      TGATACACAGCCGGGGCAA
      TCGCTTGAGTCGACTGTCCAAACATGGTTTAAGCCGCTTGCTTACGCTTTTATTCTCACAA
 40   GGGAATCTGGATACCCTCA
      GGTTTTCTACGGGGATATGTACGGGACGAAAGGAGACTCCCAGCGCGAAATTCCTGCCTTG
      AAACACAAAATTGAACCGA
      TCTTAAAAGCGAGAAAACAGTATGCGTACGGAGCACAGCATGATTATTTCGACCACCATGA
      CATTGTCGGCTGGACAAGG
 45   GAAGGCGACAGCTCGGTTGCAAATTCAGGTTTGGCGGCATTAATAACAGACGGACCCGGTG
      GGGCAAAGCGAATGTATGT
      CGGCCGGCAAAACGCCGGTGAGACATGGCATGACATTACCGGAAACCGTTCGGAGCCGGTT
      GTCATCAATTCGGAAGGCT
      GGGGAGAGTTTCACGTAAACGGCGGGTCGGTTTCAATTTATGTTCAAAGATAG
 50
```

60

5   SEQ ID No. 8:

MKQQKRLYARLLTLLFALIFLLPHSAAAAanlnapgcrvdyavtnqwpggfganvtitnlg
dpvsswkldwtytagqriq
qlwngtastnggqvsvtslpwngsiptggtasfgfngswagsnptpasfslngttctgtvp
10   ttsptrANLNGTLMQYFEW
YMPNDGQHWRRLQNDSAYLAEHGITAVWIPPAYKGTSQADVGYGAYDLYDLGEFHQKGTVR
TKYGTKGELQSAIKSLHSR
DINVYGDVVINHKGGADATEDVTAVEVDPADRNRVISGEHLIKAWTHFHFPGRGSTYSDFK
WHWYHFDGTDWDESRKLNR
15   IYKFQGKAWDWEVSNENGNYDYLMYADIDYDHPDVAAEIKRWGTWYANELQLDGFRLDAVK
HIKFSFLRDWVNHVREKTG
KEMFTVAEYWQNDLGALENYLNKTNFNHSVFDVPLHYQFHAASTQGGGYDMRKLLNGTVVS
KHPLKSVTFVDNHDTQPGQ
SLESTVQTWFKPLAYAFILTRESGYPQVFYGDMYGTKGDSQREIPALKHKIEPILKARKQY
20   AYGAQHDYFDHHDIVGWTR
EGDSSVANSGLAALITDGPGGAKRMYVGRQNAGETWHDITGNRSEPVVINSEGWGEFHVNG
GSVSIYVQRZ

SEQ ID No. 9:

```
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAC
GACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCA
CTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGT
GAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGC
CTGCAGGTCGACGCATTCCGAATACGAGGCCTGATTAATGATTACATACGCCTCCGGGTAG
TAGACCGAGCAGCCGAGCCAGTTCAGCGCCTAAAACGCCTTATACAATTAAGCAGTTAAAG
AAGTTAGAATCTACGCTTAAAAAGCTACTTAAAAATCGATCTCGCAGTCCCGATTCGCCTA
TCAAAACCAGTTTAAATCAACTGATTAAAGGTGCCGAACGAGCTATAAATGATATAACAAT
ATTAAAGCATTAATTAGAGCAATATCAGGCCGCGCACGAAAGGCAACTTAAAAAGCGAAAG
CGCTCTACTAAACAGATTACTTTTGAAAAAGGCACATCAGTATTTAAAGCCCGAATCCTTA
TTAAGCGCCGAAATCAGGCAGATAAAGCCATACAGGCAGATAGACCTCTACCTATTAAATC
GGCTTCTAGGCGCGCTCCATCTAAATGTTCTGGCTGTGGTGTACAGGGGCATAAAATTACG
CACTACCCGAATCGATAGAACTACTCATTTTTATATAGAAGTCAGAATTCATAGTGTTTTG
ATCATTTTAAATTTTTATATGGCGGGTGGTGGGCAACTCGCTTGCGCGGGCAACTCGCTTA
CCGATTACGTTAGGGCTGATATTTACGTGAAAATCGTCAAGGGATGCAAGACCAAAGTAGT
AAAACCCCGGAAGTCAACAGCATCCAAGCCCAAGTCCTTCACGGAGAAACCCCAGCGTCCA
CATCACGAGCGAAGGACCACCTCTAGGCATCGGACGCACCATCCAATTAGAAGCAGCAAAG
CGAAACAGCCCAAGAAAAAGGTCGGCCCGTCGGCCTTTTCTGCAACGCTGATCACGGGCAG
CGATCCAACCAACACCCTCCAGAGTGACTAGGGGCGGAAATTTAAAGGGATTAATTTCCAC
TCAACCACAAATCACAGTCGTCCCCGGTATTGTCCTGCAGAATGCAATTTAAACTCTTCTG
CGAATCGCTTGGATTCCCCGCCCCTAGTCGTAGAGCTTAAAGTATGTCCCTTGTCGATGCG
ATGATACACAACATATAAATACTAGCAAGGGATGCCATGCTTGGAGGATAGCAACCGACAA
CATCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCCACAGGGGGGATCCACTAGTAAC
GGCCGCCAGTGTGCTGGAAAGCGACTTGAAACGCCCCAAATGAAGTCCTCCATCCTCGCCA
GCGTCTTCGCCACGGGCGCCGTGGCTCAAAGTGGTCCGTGGCAGCAATGTGGTGGCATCGG
ATGGCAAGGATCGACCGACTGTGTGTCGGGCTACCACTGCGTCTACCAGAACGATTGGTAC
AGCCAGTGCGTGCCTGGCGCGGCGTCGACAACGCTGCAGACATCGACCACGTCCAGGCCCA
CCGCCACCAGCACCGCCCCTCCGTCGTCCACCACCTCGCCTAGCGTGGCCAGTCCTATTCG
TCGAGAGGTCTCGCAGGATCTGTTTAACCAGTTCAATCTCTTTGCACAGTATTCTGCAGCC
GCATACTGCGGAAAAAACAATGATGCCCCAGCTGGTACAAACATTACGTGCACGGGAAATG
CCTGCCCCGAGGTAGAGAAGGCGGATGCAACGTTTCTCTACTCGTTTGAAGACTCTGGAGT
GGGCGATGTCACCGGCTTCCTTGCTCTCGACAACACGAACAAATTGATCGTCCTCTCTTTC
CGTGGCTCTCGTTCCATAGAGAACTGGATCGGGAATCTTAAGTTCCTCTTGAAAAAAATAA
ATGACATTTGCTCCGGCTGCAGGGGACATGACGGCTTCACTTCGTCCTGGAGGTCTGTAGC
CGATACGTTAAGGCAGAAGGTGGAGGATGCTGTGAGGGAGCATCCCGACTATCGCGTGGTG
TTTACCGGACATAGCTTGGGTGGTGCATTGGCAACTGTTGCCGGAGCAGACCTGCGTGGAA
ATGGGTATGATATCGACGTGTTTTCATATGGCGCCCCCGAGTCGGAAACAGGGCTTTTGC
AGAATTCCTGACCGTACAGACCGGCGGAACACTCTACCGCATTACCCACACCAATGATATT
GTCCCTAGACTCCCGCCGCGCGAATTCGGTTACAGCCATTCTAGCCCAGAATACTGGATCA
AATCTGGAACCCTTGTCCCCGTCACCCGAAACGATATCGTGAAGATAGAAGGCATCGATGC
CACCGGCGGCAATAACCGGCCGAACATTCCGGATATCCCTGCGCACCTATGGTACTTCGGG
TTAATTGGGACATGTCTTTAGTGGCCGGCGCGCGGCTGGGTCGACTCTAGCGAGCTCGAGATC
TAGAGGGTGACTGACACCTGGCGGTAGACAATCAATCCATTTCGCTATAGTTAAAGGATGG
GGATGAGGGCAATTGGTTATATGATCATGTATGTAGTGGGTGTGCATAATAGTAGTGAAAT
GGAAGCCAAGTCATGTGATTGTAATCGACCGACGGAATTGAGGATATCCGGAAATACAGAC
ACCGTGAAAGCCATGGTCTTTCCTTCGTGTAGAAGACCAGACAGACAGTCCCTGATTTACC
CTTGCACAAAGCACTAGAAAATTAGCATTCCATCCTTCTCTGCTTGCTCTGCTGATATCAC
TGTCATTCAATGCATAGCCATGAGCTCATCTTAGATCCAAGCACGTAATTCCATAGCCGAG
GTCCACAGTGGAGCAGCAACATTCCCCATCATTGCTTTCCCCAGGGGCCTCCCAACGACTA
```

62

```
AATCAAGAGTATATCTCTACCGTCCAATAGATCGTCTTCGCTTCAAAATCTTTGACAATTC
CAAGAGGGTCCCCATCCATCAAACCCAGTTCAATAATAGCCGAGATGCATGGTGGAGTCAA
TTAGGCAGTATTGCTGGAATGTCGGGCCAGTTGGCCCGGGTGGTCATTGGCCGCCTGTGAT
GCCATCTGCCACTAAATCCGATCATTGATCCACCGCCCACGAGGCGCGTCTTTGCTTTTTG
```
5
```
CGCGGCGTCCAGGTTCAACTCTCTCGCTCTAGATATCGATGAATTCACTGGCCGTCGTTTT
ACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCC
CCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGC
GCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTAT
TTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAG
```
10
```
CCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCG
CTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATC
ACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATG
ATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTA
TTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATA
```
15
```
AATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTT
ATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAG
TAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAG
CGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAA
GTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCC
```
20
```
GCATACACTATTCTCAGAATGACTTGGTTGACGCGTCACCAGTCACAGAAAAGCATCTTAC
GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCG
GCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACA
TGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAA
CGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACT
```
25
```
GGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAG
TTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGG
AGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCC
CGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGA
TCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATA
```
30
```
TATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTT
TTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACC
CCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTT
GCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACT
CTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGT
```
35
```
AGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCT
AATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCA
AGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGC
CCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACA
```
40
```
GGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGT
TTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATG
GAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCAC
ATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG
CTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGA
```
45
```
AGAGAG
```

SEQ ID No. 10:

50
```
GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAC
```

63

```
GACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCA
CTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGT
GAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGC
CTGCAGGTCGACGCATTCCGAATACGAGGCCTGATTAATGATTACATACGCCTCCGGGTAG
```
5
```
TAGACCGAGCAGCCGAGCCAGTTCAGCGCCTAAAACGCCTTATACAATTAAGCAGTTAAAG
AAGTTAGAATCTACGCTTAAAAAGCTACTTAAAAATCGATCTCGCAGTCCCGATTCGCCTA
TCAAAACCAGTTTAAATCAACTGATTAAAGGTGCCGAACGAGCTATAAATGATATAACAAT
ATTAAAGCATTAATTAGAGCAATATCAGGCCGCGCACGAAAGGCAACTTAAAAAGCGAAAG
CGCTCTACTAAACAGATTACTTTTGAAAAAGGCACATCAGTATTTAAAGCCCGAATCCTTA
```
10
```
TTAAGCGCCGAAATCAGGCAGATAAAGCCATACAGGCAGATAGACCTCTACCTATTAAATC
GGCTTCTAGGCGCGCTCCATCTAAATGTTCTGGCTGTGGTGTACAGGGGCATAAAATTACG
CACTACCCGAATCGATAGAACTACTCATTTTTATATAGAAGTCAGAATTCATAGTGTTTTG
ATCATTTTAAATTTTTATATGGCGGGTGGTGGGCAACTCGCTTGCGCGGGCAACTCGCTTA
CCGATTACGTTAGGGCTGATATTTACGTGAAAATCGTCAAGGGATGCAAGACCAAAGTAGT
```
15
```
AAAACCCCGGAAGTCAACAGCATCCAAGCCCAAGTCCTTCACGGAGAAACCCCAGCGTCCA
CATCACGAGCGAAGGACCACCTCTAGGCATCGGACGCACCATCCAATTAGAAGCAGCAAAG
CGAAACAGCCCAAGAAAAAGGTCGGCCCGTCGGCCTTTTCTGCAACGCTGATCACGGGCAG
CGATCCAACCAACACCCTCCAGAGTGACTAGGGGCGGAAATTTAAAGGGATTAATTTCCAC
TCAACCACAAATCACAGTCGTCCCCGGTATTGTCCTGCAGAATGCAATTTAAACTCTTCTG
```
20
```
CGAATCGCTTGGATTCCCCGCCCCTAGTCGTAGAGCTTAAAGTATGTCCCTTGTCGATGCG
ATGATACACAACATATAAATACTAGCAAGGGATGCCATGCTTGGAGGATAGCAACCGACAA
CATCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCCACAGGGGGGATCCACTAGTAAC
GGCCGCCAGTGTGCTGGAAAGCGACTTGAAACGCCCCAAATGAAGTCCTCCATCCTCGCCA
GCGTCTTCGCCACGGGCGCCGTGGCTCAAAGTGGTCCGTGGCAGCAATGTGGTGGCATCGG
```
25
```
ATGGCAAGGATCGACCGACTGTGTGTCGGGCTACCACTGCGTCTACCAGAACGATTGGTAC
AGCCAGTGCGctagcCCTccTCGTCGAcctGTCTCGCAGGATCTGTTTAACCAGTTCAATC
TCTTTGCACAGTATTCTGCAGCCGCATACTGCGGAAAAAACAATGATGCCCCAGCTGGTAC
AAACATTACGTGCACGGGAAATGCCTGCCCCGAGGTAGAGAAGGCGGATGCAACGTTTCTC
TACTCGTTTGAAGACTCTGGAGTGGGCGATGTCACCGGCTTCCTTGCTCTCGACAACACGA
```
30
```
ACAAATTGATCGTCCTCTCTTTCCGTGGCTCTCGTTCCATAGAGAACTGGATCGGGAATCT
TAAGTTCCTCTTGAAAAAAATAAATGACATTTGCTCCGGCTGCAGGGGACATGACGGCTTC
ACTTCGTCCTGGAGGTCTGTAGCCGATACGTTAAGGCAGAAGGTGGAGGATGCTGTGAGGG
AGCATCCCGACTATCGCGTGGTGTTTACCGGACATAGCTTGGGTGGTGCATTGGCAACTGT
TGCCGGAGCAGACCTGCGTGGAAATGGGTATGATATCGACGTGTTTTCATATGGCGCCCCC
```
35
```
CGAGTCGGAAACAGGGCTTTTGCAGAATTCCTGACCGTACAGACCGGCGGAACACTCTACC
GCATTACCCACACCAATGATATTGTCCCTAGACTCCCGCCGCGCGAATTCGGTTACAGCCA
TTCTAGCCCAGAATACTGGATCAAATCTGGAACCCTTGTCCCCGTCACCCGAAACGATATC
GTGAAGATAGAAGGCATCGATGCCACCGGCGGCAATAACCGGCCGAACATTCCGGATATCC
CTGCGCACCTATGGTACTTCGGGTTAATTGGGACATGTCTTTAGTGGCCGGCGCGGCTGGG
```
40
```
TCGACTCTAGCGAGCTCGAGATCTAGAGGGTGACTGACACCTGGCGGTAGACAATCAATCC
ATTTCGCTATAGTTAAAGGATGGGGATGAGGGCAATTGGTTATATGATCATGTATGTAGTG
GGTGTGCATAATAGTAGTGAAATGGAAGCCAAGTCATGTGATTGTAATCGACCGACGGAAT
TGAGGATATCCGGAAATACAGACACCGTGAAAGCCATGGTCTTTCCTTCGTGTAGAAGACC
AGACAGACAGTCCCTGATTTACCCTTGCACAAAGCACTAGAAAATTAGCATTCCATCCTTC
```
45
```
TCTGCTTGCTCTGCTGATATCACTGTCATTCAATGCATAGCCATGAGCTCATCTTAGATCC
AAGCACGTAATTCCATAGCCGAGGTCCACAGTGGAGCAGCAACATTCCCCATCATTGCTTT
CCCCAGGGGCCTCCCAACGACTAAATCAAGAGTATATCTACCGTCCAATAGATCGTCTT
CGCTTCAAAATCTTTGACAATTCCAAGAGGGTCCCCATCCATCAAACCCAGTTCAATAATA
GCCGAGATGCATGGTGGAGTCAATTAGGCAGTATTGCTGGAATGTCGGGCCAGTTGGCCCG
```
50
```
GGTGGTCATTGGCCGCCTGTGATGCCATCTGCCACTAAATCCGATCATTGATCCACCGCCC
ACGAGGCGCGTCTTTGCTTTTTGCGCGGCGTCCAGGTTCAACTCTCTCGCTCTAGATATCG
ATGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAAC
```

64

TTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTT
CTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCT
CTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACG
5   GGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATG
TGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCC
TATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCG
GGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCG
CTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTA
10   TTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGC
TCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGT
TACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTT
TTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGC
CGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGACGCGTCA
15   CCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCA
TAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGA
GCTAACCGCTTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCG
GAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAA
CAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAAT
20   AGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGC
TGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCAC
TGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC
TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAA
CTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTA
25   AAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTT
TTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTT
TTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTT
TGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGAT
ACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCA
30   CCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGT
CGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTG
AACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATAC
CTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATC
CGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTG
35   GTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGC
TCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG
CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAA
CCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGC
GAGTCAGTGAGCGAGGAAGCGGAAGAGAG
40

SEQ ID No. 11:

GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAC
45   GACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCA
CTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGT
GAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGC
CTGCAGGTCGACGCATTCCGAATACGAGGCCTGATTAATGATTACATACGCCTCCGGGTAG
TAGACCGAGCAGCCGAGCCAGTTCAGCGCCTAAAACGCCTTATACAATTAAGCAGTTAAAG
50   AAGTTAGAATCTACGCTTAAAAAGCTACTTAAAAATCGATCTCGCAGTCCCGATTCGCCTA
TCAAAACCAGTTTAAATCAACTGATTAAAGGTGCCGAACGAGCTATAAATGATATAACAAT
ATTAAAGCATTAATTAGAGCAATATCAGGCCGCGCACGAAAGGCAACTTAAAAAGCGAAAG

65

```
CGCTCTACTAAACAGATTACTTTTGAAAAAGGCACATCAGTATTTAAAGCCCGAATCCTTA
TTAAGCGCCGAAATCAGGCAGATAAAGCCATACAGGCAGATAGACCTCTACCTATTAAATC
GGCTTCTAGGCGCGCTCCATCTAAATGTTCTGGCTGTGGTGTACAGGGGCATAAAATTACG
CACTACCCGAATCGATAGAACTACTCATTTTTATATAGAAGTCAGAATTCATAGTGTTTTG
ATCATTTTAAATTTTTATATGGCGGGTGGTGGGCAACTCGCTTGCGCGGGCAACTCGCTTA
CCGATTACGTTAGGGCTGATATTTACGTGAAAATCGTCAAGGGATGCAAGACCAAAGTAGT
AAAACCCCGGAAGTCAACAGCATCCAAGCCCAAGTCCTTCACGGAGAAACCCCAGCGTCCA
CATCACGAGCGAAGGACCACCTCTAGGCATCGGACGCACCATCCAATTAGAAGCAGCAAAG
CGAAACAGCCCAAGAAAAAGGTCGGCCCGTCGGCCTTTTCTGCAACGCTGATCACGGGCAG
TCAACCACAAATCACAGTCGTCCCCGGTATTGTCCTGCAGAATGCAATTTAAAACTCTTCTG
CGAATCGCTTGGATTCCCCGCCCCTAGTCGTAGAGCTTAAAGTATGTCCCTTGTCGATGCG
ATGATACACAACATATAAATACTAGCAAGGGATGCCATGCTTGGAGGATAGCAACCGACAA
CATCACATCAAGCTCTCCCTTCTCTGAACAATAAACCCCACAGGGGGGATCCACTAGTAAC
GGCCGCCAGTGTGCTGGAAAGCGACTTGAAACGCCCCAAATGAAGTCCTCCATCCTCGCCA
GCGTCTTCGCCACGGGCGCCGTGGCTCAAAGTGGTCCGTGGCAGCAATGTGGTGGCATCGG
ATGGCAAGGATCGACCGACTGTGTGTCGGGCTACCACTGCGTCTACCAGAACGATTGGTAC
AGCCAGTGCGCTAGCGTCCAGATCCCCTCCAGCAGCACCAGCTCTACCAGAACGATTGGTAC
CCAGCACCAGCACCACGTCCACCTCCACCACCTCGAGCCCGCCAGTCCAGCCTACGACTCC
CAGCGCTAGCCCTCCTCGTCGACCTGTCTCGCAGGATCTGTTTAACCAGTTCAATCTCTTT
GCACAGTATTCTGCAGCCGCATACTGCGGAAAAAACAATGATGCCCCAGCTGGTACAAACA
TTACGTGCACGGGAAATGCCTGCCCCGAGGTAGAGAAGGCGGATGCAACGTTTCTCTACTC
GTTTGAAGACTCTGGAGTGGGCGATGTCACCGGCTTCCTTGCTCTCGACAACACGAACAAA
TTGATCGTCCTCTCTTTCCGTGGCTCTCGTTCCATAGAGAACTGGATCGGGAATCTTAAGT
TCCTCTTGAAAAAAATAAATGACATTTGCTCCGGCTGCAGGGGACATGACGGCTTCACTTC
GTCCTGGAGGTCTGTAGCCGATACGTTAAGGCAGAAGGTGGAGGATGCTGTGAGGGAGCAT
CCCGACTATCGCGTGGTGTTTACCGGACATAGCTTGGGTGGTGCATTGGCAACTGTTGCCG
GAGCAGACCTGCGTGGAAATGGGTATGATATCGACGTGTTTTCATATGGCGCCCCCGAGT
CGGAAACAGGGCTTTTGCAGAATTCCTGACCGTACAGACCGGCGGAACACTCTACCGCATT
ACCCACACCAATGATATTGTCCCTAGACTCCCGCCGCGCGAATTCGGTTACAGCCATTCTA
GCCCAGAATACTGGATCAAATCTGGAACCCTTGTCCCCGTCACCCGAAACGATATCGTGAA
GATAGAAGGCATCGATGCCACCGGCGGCAATAACCGGCCGAACATTCCGGATATCCCTGCG
CACCTATGGTACTTCGGGTTAATTGGGACATGTCTTTAGTGGCCGGCGCGGCTGGGTCGAC
TCTAGCGAGCTCGAGATCTAGAGGGTGACTGACACCTGGCGGTAGACAATCAATCCATTTC
GCTATAGTTAAAGGATGGGGATGAGGGCAATTGGTTATATGATCATGTATGTAGTGGGTGT
GCATAATAGTAGTGAAATGGAAGCCAAGTCATGTGATTGTAATCGACCGACGGAATTGAGG
ATATCCGGAAATACAGACACCGTGAAAGCCATGGTCTTTCCTTCGTGTAGAAGACCAGACA
GACAGTCCCTGATTTACCCTTGCACAAAGCACTAGAAAATTAGCATTCCATCCTTCTCTGC
TTGCTCTGCTGATATCACTGTCATTCAATGCATAGCCATGAGCTCATCTTAGATCCAAGCA
CGTAATTCCATAGCCGAGGTCCACAGTGGAGCAGCAACATTCCCCATCATTGCTTTCCCCA
GGGGCCTCCCAACGACTAAATCAAGAGTATATCTCTACCGTCCAATAGATCGTCTTCGCTT
CAAAATCTTTGACAATTCCAAGAGGGTCCCCATCCATCAAACCCAGTTCAATAATAGCCGA
GATGCATGGTGGAGTCAATTAGGCAGTATTGCTGGAATGTCGGGCCAGTTGGCCCGGGTGG
TCATTGGCCGCCTGTGATGCCATCTGCCACTAAATCCGATCATTGATCCACCGCCCACGAG
GCGCGTCTTTGCTTTTTGCGCGGCGTCCAGGTTCAACTCTCTCGCTCTAGATATCGATGAA
TTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAAT
CGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATC
GCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCT
TACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGAT
GCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTT
GTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCA
GAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTT
```

66

```
TTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAA
ATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCAT
GAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAA
CATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACC
CAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACAT
CGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCA
ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGC
AAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGACGCGTCACCAGT
CACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACC
ATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAA
CCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCT
GAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACG
TTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACT
GGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTT
TATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGG
CCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGG
ATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTC
AGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGG
ATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGT
TCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCT
GCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCG
GATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAA
ATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCC
TACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGT
CTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGG
GGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACA
GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTA
AGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATC
TTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTC
AGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTT
TTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA
TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTC
AGTGAGCGAGGAAGCGGAAGAGAG
```

67

CLAIMS

1. A process for desizing cellulose-containing fabric or tex-
tile, wherein said fabric or textile is treated with a
5 modified enzyme (enzyme hybrid) which comprises a
catalytically active amino acid sequence of a non-cellulolytic
enzyme linked to an amino acid sequence comprising a
cellulose-binding domain.

10 2. A process according to claim 1, wherein said catalytically
active amino acid sequence derives from an enzyme selected
from the group consisting of amylases and lipases.

3. A process according to claim 2, wherein said amylase is an
15 α-amylase obtainable from a species of *Bacillus*.

4. A process according to claim 2 or 3, wherein said α-amylase
is obtainable from *Bacillus licheniformis*.

20 5. A process according to any one of claims 2-4, wherein an
amylolytic enzyme hybrid is employed in an amount
corresponding to an amylase activity in the range of between 1
and 5000 KNU per litre of desizing liquor.

25 6. A process according to claim 2, wherein said lipase is
obtainable from a species of *Humicola*, *Candida*, *Pseudomonas* or
Bacillus.

7. A process according to claim 2 or 6, wherein a lipolytic
30 enzyme hybrid is employed in an amount corresponding to a
lipase activity in the range of between 10 and 20000 LU per
litre of desizing liquor.

8. A process according to claim 1, wherein said cellulose-

binding domain is obtainable from a cellulase, a xylanase, a mannanase, an arabinofuranosidase, an acetylesterase or a chitinase.

5 9. A process according to claim 1, wherein said enzyme hybrid is obtained by a method comprising growing a transformed host cell containing an expression cassette which comprises a DNA sequence encoding said enzyme hybrid, whereby said enzyme hybrid is expressed.

10

10. A desizing composition comprising:

an enzyme hybrid which comprises a catalytically active amino acid sequence of a non-cellulolytic enzyme linked to an amino
15 acid sequence comprising a cellulose-binding domain; and

a wetting agent.

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C12N 9/00, D06M 16/00, C07K 19/00 // C11D003386

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EDOC, BIOSIS, DBA, CA, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 9311249 A1 (NOVO NORDISK A/S), 10 June 1993 (10.06.93) --	1-10
A	WO 9305226 A1 (UNIVERSITY OF BRITISH COLUMBIA), 18 March 1993 (18.03.93) -- --------	1-10

☐ Further documents are listed in the continuation of Box C.   ☒ See patent family annex.

Date of the actual completion of the international search	Date of mailing of the international search report
6 May 1997	22 -05- 1997
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Patrick Andersson Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1992)

Patent document cited in search report			Publication date	Patent family member(s)			Publication date
WO	9311249	A1	10/06/93	BR EP FI JP	9206866 0618974 942644 8504560	A A A T	21/11/95 12/10/94 03/06/94 21/05/96
WO	9305226	A1	18/03/93	NONE			